

Promises and pitfalls for new year

From what has been learned in 1982, the year ahead will be full of excitement. And in spite of the recession, most governments plan to protect basic research. But there are pitfalls to be avoided.

THE scientific enterprise is in remarkably good shape. That is the unexpected truth at the beginning of a new year, and after almost a decade of discontent in the scientific profession about the effects of public parsimony on the fruitfulness of research. Two kinds of things have happened. First, governments of very different colours have woken up to the need that they should provide generous support for basic research. Second, and more important, the pace of discovery has continued unchecked in a host of different directions. The result is that 1983 will be a year in which people can confidently expect rapid progress towards the understanding of several long-standing conundrums, among which the mechanism of carcinogenesis is merely the most striking. Before the new year is out, there will be a rich harvest of discoveries to celebrate. What will have been, for many people, a respite from the endless task of reading the journals is a time for taking stock.

That 1983 may be the year in which carcinogenesis is finally understood is obviously arresting, but that prospect is also a telling pointer to the nature and texture of discovery. Even since the appearance thirty years ago of the paper on the structure of DNA by Watson and Crick (*Nature* will be holding a conference to mark that anniversary in April), it has been natural that people should have looked for an explanation of carcinogenesis in terms of alterations of the structure of normal genetic material. Quite early on, it was established that certain planar aromatic molecules which are also carcinogenic, anthraquinone for example, can physically interpolate themselves between successive turns of a DNA helix, suggesting that faithful DNA replication might by this means be prevented. Unfortunately, too little was known, thirty years ago, of the complexity of normal genetic material; genes are not in general simply connected stretches of DNA but may be interrupted by regions called introns whose function is still unclear. And at that early stage, even the mechanism by which the activity of genes is regulated in the simplest bacteria was unknown; the search for an understanding of the more complicated processes by which this is accomplished in the cells of higher organisms, those capable of contracting cancer, still continues.

Cancer

So, not unexpectedly, early expectations that the simple intercalation of appropriate molecules in genetic DNA might account for the causation of cancer were disappointed. Maybe such processes are involved in some kinds of cancers, certainly some carcinogens appear to cause specific and possibly significant point mutations in DNA molecules, but there can be no such simple general rule for the causation of cancer. So much is clear from what has been learned, for example, of the biochemical transformation of materials such as benzpyrene in normal cells.

So what reason is there to expect that recent discoveries are more significant? In the past eighteen months, it has been established by energetic groups in the United States that certain kinds of tumour cells are characterized by the possession within their genetic complement of genes that can transform certain other cells (specifically those of the strain of mouse cells called NIH 3T3) into frankly malignant condition; that these oncogenes are related (strictly speaking, are homologous) to the oncogenes of the RNA viruses known as retroviruses, themselves linked with

the occurrence of many animal cancers; that cells of one specific human cancer (a bladder cancer) contain a gene that differs from the normal version by the replacement of only one nucleotide (guanine) by another (thymine); and that the gene concerned is normally responsible for the production of an essential protein called p21 whose function is, nevertheless, unknown. While nobody in his senses would expect that these discoveries point to a unique causation of all cancers, what makes them important is their concatenation. Almost for the first time, there is a group of phenomena still fully to be explored that must have a bearing on the origin of a particular kind of cancer. The richness of the diversity of these phenomena is the reason why, for the first time, there is a chance of getting to the bottom of the phenomenon of cancer.

Prevention

Naturally enough, excitement at the immediate prospect has bounded ahead. Part of the explanation is that people now know that the techniques exist for characterizing oncogenes other than that fully described for a line of cells from a single human bladder carcinoma; it is astonishing that merely thirty years after Watson and Crick's well-educated guess about the structure of DNA should have been universally recognized as true, it is possible to tell how the nucleotides are arranged in a naturally occurring molecule at a rate in excess of 1,000 bases a day and, even more to the point, to fish out from an entire genome those parts of the functional DNA likely to have some particular function. The unsung heroes of the forthcoming understanding of malignancy (nevertheless not noticeably deficient in Nobel medals) are those who have made child's play of telling how DNA is put together. There can never have been a technique of what is essentially chemical analysis that is more potent and potentially more pregnant than the technology that skilled hands now have within reach for telling how genetic material is constructed. What would Mendel have thought of that?

The more important part of the explanation is the glitter of the prize to be won. In advanced societies, cancer of one kind or another kills a third of those who die. Nobody pretends that an understanding of the mechanisms of carcinogenesis will imply a cure, or that remarkably earlier diagnosis makes possible markedly better treatment. But there is at least a chance that better understanding will reinforce what is known empirically about prevention ("Do not smoke!") and a certainty that understanding what is really wrong with malignant cells will so illuminate the general understanding of cell biology — ordinary people's understanding that they are the products of their cells and of their cell's progeny — that ordinary people will be able to behave more rationally. Nobody now alive should confidently expect to be cured of a cancer yet to be contracted by what is about to be discovered about the mechanisms of malignancy; there is a chance.

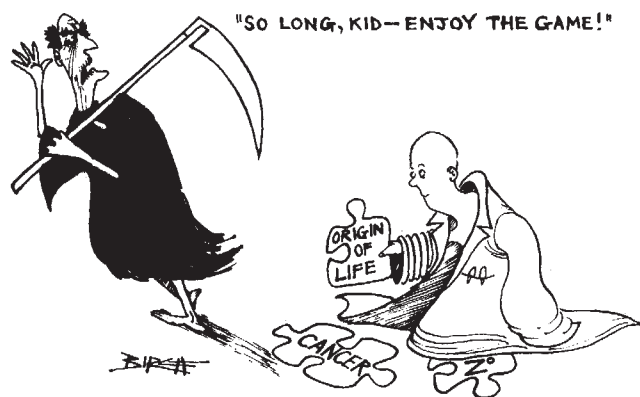
Unfashionable though that may be, much the same is true of the impending test of the reality (or otherwise) of the intermediate vector boson. Whichever way that question is resolved (it exists or it does not, within the capacity of machines such as the CERN proton-antiproton collider at Geneva), the outcome will be electrifying. Success will mean that Weinberg and Salam have been right all this decade past in their physical intuition that of the



four kinds of forces that can exist between particles of matter, electromagnetic forces and weak nuclear forces (which bind electrons with protons to make neutrons or, more generally, leptons with hadrons) have more in common with each other than either can have with gravitational forces on the one hand and strong nuclear forces on the other. Failure will mean nothing less than a reconsideration of the conventional wisdom that the interactions between elementary particles (whatever *they* are) come in only four forms. Nobody at this stage will pretend that a resolution of that question will accomplish much in the real world — bring the long recession to an end, for example. But nobody can claim that crucial test of theories of how matter is constructed can end in frustration without throwing everything (and everybody) into confusion. *Nature's* bet, for what it is worth, is that during 1983 the intermediate vector boson will be found to exist but also to be more massive than the equivalent of 80 GeV conventionally predicted. The appearances, in other words, will be confirmed.

Understanding

It would be proper at this stage to protest that these issues — what causes cancer and what are the interactions between different kinds of matter — are not central to science in 1983. In one sense, that complaint is profoundly mistaken. In the evolution of scientific understanding ("progress" is a more appropriate but more pejorative term), the acid tests of credibility are those which apply to the most radical predictions. If a man should say that the perihelion of Mercury must on general considerations precess by 0.42 seconds of arc a century and be proved right by observation, the integrity of the scientific enterprise must be confirmed not merely in the eyes of other professional people, would-be recruits to the scientific profession and of custodians of public purses, but also among the general public. If a man (or several people) should chance his arm by saying that there must be a black hole at the centre of every other galaxy, be professionally acclaimed for having done so and then be proved wrong, the profession will be set back on its heels. The moral, for those who feel the wind blowing in their sails during 1983, should be to strive towards sobriety. The molecular biology of carcinogenesis is ripe for picking, the intermediate vector boson perhaps dangerously over-ripe. Elsewhere, the need is that, with the wind at their backs, professional people should be professionally discreet.



During 1983, there will be no shortage of less spectacular discoveries and opportunities. Extraterrestrially, the new millisecond pulsar, undoubtedly only the first of its kind to be discovered, is a sign that neutron stars can be made by mechanisms other than supernova explosions. A little reflection will show that even in the supposedly well-known galaxy to which the Sun belongs, new kinds of celestial objects are being discovered more frequently now than half a century ago. As what passes in these spectacular days almost as a matter of routine, however, the millisecond pulsar is a telling reminder of the importance of technique. The object has been known as a radio source for twenty years: only in the past year or so have computers

been commodious enough to sample, say, a ten minute record of the output of the source at intervals of, say, a microsecond.

As things have turned out, it will be disappointing if there are not half a dozen similar sources by the end of the year. With luck, the Very Large Array of radiotelescopes in New Mexico should also, during the year, make possible the direct observation of the process of star formation in some of the relatively dense molecular clouds in nebulae such as that in Orion. The more detailed description of the motion of visible objects within the galaxy will however wait on the launching of the large space telescope, at the soonest three years away. Meanwhile, speculation about the galactic centre (like the search for extra-terrestrial life to which the US National Aeronautics and Space Administration seems recently to have taken a liking) should be discouraged. Extragalactically, the crucial question to be decided is how (and when) galaxies were formed. The large space telescope should help again, as will conventional observation; speculation about the first few fractions of a femtosecond after the supposed Big Bang should wait on a resolution of crucial questions in high-energy physics. The chances that there will be during 1983 a convincing start on the problem of how to reconcile quantum mechanics and gravitation, increasingly an impediment to a proper understanding of energetic phenomena in the universe at large, are small; the spadework done so far serves chiefly to emphasize not merely the technical but the conceptual difficulties of the problem.

Enlivenment

Nearer home, within the Solar System, 1983 is more likely to be disappointing than enlivening. Planetary scientists complain that they are being starved of new observational data with which to make sense of the relationship between the Sun and its planets and their satellites. In reality, however, the data now available exceed in their refinement by several orders of magnitude those available, say, twenty years ago, yet the understanding of how the Solar System was formed has hardly advanced at all. Is it time that planetary scientists were required to travel less and to think more? And might the same be true of the geophysicists who seem (with general acclaim) to have given full-throated approval to the doctrine of plate tectonics without being able to give a simple explanation of the consequences (for Asia) of the northward movement of India or of the origin of the western seaboard of North America? Curiously enough, one of the best bets for the year ahead is that there will be a better understanding of what happens within the Sun and near its surface, partly from the continuing observation of the solar wind but primarily because of a deeper conceptual understanding of the structure of a rotating middle-sized middle-aged star, with all the consequences that may have for the understanding of past climatic variations on the surface of the Earth. (Terrestrial climatic modelling is, however, likely to remain trapped in that version of Parkinson's law which states that the complexity of any problem will increase so as to over-tax whatever computer power is available.)

The origin of life on the surface of the Earth is less of a problem, or is headed in that direction. It seems to be agreed that it will not in the near future be possible to tell just how living things came to live, but that it will be possible to demonstrate that chemical processes can be a sufficient explanation. RNA molecules are the preferred progenitors of self-replicating entities, and the search for plausible catalysts may be enlivened by the discovery during 1982 that naturally occurring RNA molecules may sometimes serve as enzymes for their own transformation. The implications of such processes for the understanding of the thermodynamics of systems (biological or otherwise) far from a state of equilibrium are, however, unclear. A small army of theoreticians could be well occupied, during the year ahead, in coming to grips with the energetics of processes as different as photosynthesis, the action of muscle fibres and the behaviour of the Earth's atmosphere and oceans (which likewise are not in equilibrium in the strict sense but whose behaviour is driven by energy fluxes, mostly from the Sun). And, in passing, is it too much to hope that during the year ahead, somebody somewhere will put an end to the continuing puzzle

about ball-lightning, the well-attested phenomenon still lacking an explanation?

All these, of course, are topics in which discovery lifts public esteem for the scientific enterprise. (Sadly, discoveries in mathematics such as a new general solution for a class of non-linear differential equations describing almost physical entities called solitons, or classifications of the ways in which three-dimensional space may be closed, go almost unnoticed even among professionals.) Equally important, in that regard, are the applications now flooding from the laboratories. What will microchips be called when their characteristic dimensions are measured in a few tens and not a few thousands of interatomic spacings? The year ahead should at least give some clues to what these devices will make possible. And what will become of chemical technology when the trick is not that of designing a process plant that will enable chemicals to interact by the tonne but that of synthesizing a few kilograms of an appropriate catalyst instead? That ambition, which has repeatedly waxed and waned since the 1890s, is now again waxing. If chlorophyll, a kind of semiconductor molecule, will in the presence of sunlight turn carbon dioxide into sugar, why not devise some molecular semiconductor that will turn sugar into, say, steroids (or even just polyethylene)?

Biotechnology

Speculative technology such as this should not turn attention away from the sure success for 1983 — that of biotechnology. It is easy to complain that hitherto biotechnology has put on the market a vaccine against a disease of pigs and an authentic human insulin that can, in any case, be manufactured chemically from animal insulin. The big prizes lie in the future — the near future. Biologically engineered proteins of the human blood plasma will abound during 1983. So will lymphokines, including the interferons and interleukins which may find a place in the clinical modulation of the immune system. Specific agents active in the central nervous system will follow. The steady deepening of understanding of the mammalian immune system in the past few years, which has already provided a more or less convincing map of where the genes lie, must soon yield practical benefits, certain to be multiplied when the histocompatibility genes are as well understood. Before then, however, engineered vaccines against infections are different as influenza and malaria will be on the market. And gene transplantation, not so far attempted in the treatment of human deficiency diseases for perfectly good reasons (ethical as well as technical) has already made abnormally large mice. Will larger and perhaps more productive farm animals be the next step?

Prizes

All this, and the understanding of carcinogenesis as well? It is no wonder that governments around the world have bought the bill of goods on offer. Separately, their record in the past year or so is quite remarkable. The signs are that President Reagan's third budget, due this month, will be more generous towards the government agencies making research grants (chiefly the National Science Foundation and the National Institutes of Health) than optimistic realists had been hoping. The West German government, inheriting an unfamiliar economic crisis and faced with a general election in only two months, has nevertheless declared that basic research will be protected. For the past year, the government of France has promised even more (but may yet be caught out by the inflation its general policies threaten). It is too soon to know how far the new government in Spain will go along that road, but it is clear that the British government has honoured Mrs Thatcher's promise three years ago that what is called the science budget would be protected — the snag is merely that the capacity of academic researchers to use resources effectively has been undermined by the same government's short-sighted policy towards the universities. Even so, and even in Britain, the largely unprompted conversion of Western governments to the long-term importance of basic research must be welcomed. That it springs from a deep anxiety about the economic future of the

industrialized world is, in the short run, neither here nor there.

On the face of things, then, 1983 will be a year of great achievement and nobody will have reason to complain. There are, however, two important pitfalls in the compacts that Western governments are offering to their research communities. For technical reasons, these governments will channel a large part of their largesse through their own well-established research institutions, whose functions are far from clear. In the United States, for example, there are two nuclear weapons laboratories when one (at the most) would suffice, two (Argonne and Oak Ridge) conceived of after the Second World War as places for the support of civilian nuclear power whose roles have never been redefined and four high-energy physics laboratories (not counting Cornell). Are these all necessary? The commission set up a year ago by Dr George ("Jay") Keyworth to examine the functions of these laboratories, and which is soon due to report, had better ask itself how far the government's welcome commitment to basic research is blunted by the continued dependence of these institutions on public funds.

The standard compromise, now being urged in West Germany on the dozen or so national laboratories, is that such institutions should not be drastically changed but, rather, encouraged to work to more immediate goals, of greater benefit to industry. The trouble, as the experience of British laboratories such as Harwell in the 1970s shows, is that while an imaginative directorate can raise a large fraction of the payroll by means of industrial contracts, industry will rightly prevent such a laboratory from being a prime mover. If governments wish their belief in basic research to pay off, they must seriously redirect what they spend towards independent researchers and away from their own in-house laboratories.

Snags

The second important snag in the prospectus now on offer is the unashamed declaration of the governments concerned that they want prosperity to follow from what they spend. In the free-market economies of the United States and West Germany, partnerships between industrial corporations and basic researchers are to be encouraged by fiscal means (but the West German government has not yet said how). In Britain, the research councils and the central government are doing everything they can to encourage the industrial application of research (another £10 million for information technology was found only the week before Christmas). Socialist France is more convinced of the benefits of central planning (which worked well enough twenty years ago, in Gaullist times). The twin dangers in these policies are that all the interested governments will plan for the same obvious commercial benefits from basic research — and that all but one of them will be disappointed — and that their attention spans are electorally limited.

This is not, after all, the first time that the same governments have blessed basic research — and then turned cool. Immediately after the Second World War, basic research was almost everywhere the putative harbinger of prosperity. Then too, if relatively, the intellectual prospect was as exciting as now. (Vannevar Bush's eloquent document, "The Endless Frontier", refers.) So what went wrong? Not basic research. Indeed, there is a case for saying that the prosperity of the 1960s derived from the foundations laid in the late 1940s. But governments everywhere turned sour because they found that even the prosperity of that now seemingly marvellous decade of economic growth was not limitless, as they chose to think. This time, please, let modesty prevail. To understand what causes cancer, or to make a machine that will think, is worthwhile and within reach. To tell where these ends will lead is in the nature of things not possible; to know that the consequences of discovery are usually beneficial is, for most people, sufficient. But what of the other side of that coin? May not the spate of discovery now in prospect be subversive of the present scheme of things? That, in the past decade or so, has been a constant anxiety. It is mistaken. Past changes have improved our lives. Present changes will be as welcome at least to those who survive. That is what new years are about.

Listen to Andropov

Mr Yuri Andropov's views on arms control, unacceptable as yet, deserve close attention.

ARMS control negotiations are necessarily like poker games. Even when the participants are more or less agreed on what would make an acceptable agreement (which rarely happens), they keep significant information up their sleeves and separately seek to push the basis of the negotiations in a direction that will give them some advantage, military or political. Mr Yuri Andropov's public statements in the past few weeks on the conduct of the parallel arms control negotiations at Geneva, in themselves a breach of the original agreement with the United States that the negotiations would be kept confidential, are part of the same poker game. By presenting himself to the world at large, but especially to the electorates of Western Europe, as a man of peace whose patience with the slow progress at Geneva has been exhausted, Mr Andropov will win friends both for himself and for the Soviet Union. Yet what he has to say is too interesting to be quickly dismissed. Governments in the West were in particular too quick to denounce Andropov's mid-December proposals on intermediate range nuclear missiles as unacceptable. They are that, but they are also a much better basis for negotiation than seemed likely only a few months ago.

The complexity of the Geneva process should not be underestimated. There are two ostensibly independent sets of negotiations under way, on nuclear weapons of intermediate range and on strategic weapons. Both negotiations are bilateral, between the United States and the Soviet Union, but the principals consult with their allies from time to time. The first set of negotiations is partly stimulated by the decision of NATO governments in 1979 that they would install intermediate range nuclear weapons in Europe beginning at the end of 1983 if before then they had not secured an acceptable agreement on the reduction of Soviet nuclear forces. The negotiations on strategic arms, on the other hand, are a continuation of previous bilateral negotiations, especially of the Strategic Arms Limitation Talks (SALT), made necessary by the undertakings the superpowers have given to others (for example, the members of the Non-Proliferation Treaty), by their own anxieties about the dangers (and the costs) of the strategic arms race — and made possible by President Reagan's belated acknowledgement, towards the end of his first year in office, that arms control negotiations were unavoidable. The two sets of negotiations are being conducted separately because they have different origins, but they are technically so closely linked that they must sooner or later be thrown together. It is, for example, unthinkable that the Soviet Union would settle for an agreed reduction of strategic missiles if other missiles capable of serving strategic purposes were simultaneously being installed in Western Europe. Mr Andropov, having broken his predecessor's vow of silence, might usefully have acknowledged that the Geneva process must eventually be all of a piece.

Although Andropov's proposals on intermediate weapons are, as described, unacceptable to the West, this does not imply that they are without merit. What he suggests is that the Soviet missiles which principally threaten Western European targets, the mobile SS20s, should be withdrawn out of range except for a force equivalent (in some sense) to the combined strengths of the British and French nuclear forces. In return, the United States would not go ahead with the supply of cruise and Pershing II missiles for bases in Western Europe.

The weakness of this proposal is that it would still leave in Soviet hands a dominant intermediate nuclear striking force represented by the substantial force of aircraft based in the Soviet Union. Moreover, the SS20 missiles, being mobile, could quickly be redeployed so as to threaten Europe, which is why the West at Geneva has been looking for verifiable destruction of missile stocks. The public Soviet admission that its large force of SS20 missiles is a cause for anxiety in Western Europe is nevertheless

important. While this point has been emphasized by the United States behind the closed doors at Geneva, a public admission that it may be valid is worth seizing. Andropov's suggestion that the SS20 force might be reduced to match the French and British nuclear forces is more of a red herring. Both the British and French governments are mistaken in their assumption so far that their nuclear forces are irrelevant to the Geneva talks, but Andropov is wrong simply to lump them in with the intermediate nuclear weapons. Those carried by submarines, and whose accuracy is inherently less than that of land-based weapons, belong in the strategic talks under way at Geneva (which is another reason for throwing the two processes together).

Andropov's proposals on the reduction of strategic arms are more realistic. Limiting the numbers of strategic missiles held by the superpowers to 1,800 each would imply a reduction of about a quarter. Each side would be left with some 5,000 nuclear warheads, sufficient both for destroying the other in a final act of retaliation and, more significantly, for planning a first strike against the other's missile forces. Eventually, the objective in the SALT talks should be such a reduction of strategic forces that a first-strike attack is not feasible. It is probably too much to hope that either of the governments concerned would be able to go that far when they have so little recent experience to show that they can trust each other. Living with a 25 per cent reduction for a few years would moreover provide an opportunity for testing the efficacy of the procedures of verification on which the United States Senate is certain to spend much of its energy if some agreement eventually emerges from Geneva. To its credit, the United States Administration has in the past few days been more welcoming towards this Andropov proposal than towards what he had to say about intermediate weapons. The overriding snag is that the two sets of negotiations will have to be brought to a conclusion more or less together.

What of development?

The Brandt Commission's second report should reawaken compassion for the poor.

As the long recession continues into another year, the prospects worsen for those developing countries which have for decades been the poorest in the world. Their capacity to fend for themselves has for the past five years been undermined by the slump in the prices of primary commodities, likely to continue while the recession lasts. The past few months have brought two further threats to their economic survival — the rash of protectionism that has affected industrialized countries, more concerned with the preservation of domestic jobs than with the capacity of the developing countries to eke out a living by the export of simple materials such as shoes and textiles, and by fears of instability in the international banking system. The poorest countries are not, of course, among the principal debtors of the banking system, which has principally been lending to countries such as Brazil and Mexico which seemed to have a chance, a few years ago, of breaking through to industrial prosperity. But now that the banking system has taken fright, the poorest countries will be less able than ever to hope that a commercial loan will help tide them over a crisis.

The recession apart, the steady deterioration of the poorest countries will bring trouble for the prosperous world. The need now is that the obligations of the industrialized world towards the poorest countries should be more widely understood. There will be an opportunity later this month, when the Brandt Commission is due to publish a second instalment of its continuing study of the problems of development. Although the commission is rightly concerned, at this stage, with difficulties such as those that have arisen in the banking system, its new report should rekindle interest in a daunting but important problem. While, in the short run, the plight of the poorest countries may be eased by the creation of more credit, ultimately there must be no escape from a more substantial transfer of resources from the prosperous to the needy.

US cancer research

Oncogenes cause cancer institute to change tack

Washington

Now that some specific human genes have been implicated in the causation of cancer, what should happen to more traditional areas of cancer research? That question is being asked by policymakers at the National Cancer Institute (NCI) faced with the task of distributing \$983 million a year on cancer research.

Since the discovery that a difference of just one nucleotide can distinguish a human cancer gene from its normal counterpart (see *Nature* 11 November 1982, p.103), interest in the study of oncogenes has increased dramatically. Some researchers in the field are hoping for a corresponding increase in NCI's involvement with molecular and cellular genetics. But NCI's policy at the moment is, in the words of its director, Vincent DeVita Jr, "to let the system drive itself". That is, he wants researchers at the "cutting edge" of the field, and those in affected areas, to initiate proposals and have the system review them, so the work will expand only as the peer review system allows.

Grant applications in NCI's \$100 million competing grants programme are scored by their peer reviewers. DeVita says that proposals relating to oncogenes and sequencing have been getting the highest scores lately, so the work seems to be expanding naturally. DeVita says his policy choice is not whether to expand the attention given to oncogenes, but whether artificially to set limits on it, and this he declines to do.

One exception to NCI's *laissez faire* policy, however, is the attempt to make the Frederick cancer research facility in Frederick, Maryland into a major centre of activity of the "cutting edge". DeVita is looking for a new principal investigator to run Frederick's basic research programme, and is hoping to get someone of international stature in the field of oncogenes.

The new principal investigator at Frederick will manage not only the \$8 million basic research programme there, but also an additional \$2 million to be re-programmed from other work.

DeVita has already enhanced Frederick's role in key areas of molecular genetics. Two years ago he invited some of the groups at NCI to move from Bethesda out to Frederick, which they did. And a year ago DeVita appointed Peter Fischinger, who worked with George Todaro in NCI's viral carcinogenesis laboratory, to be associate director of NCI. Frederick is one of his responsibilities.

Another is keeping track of the emerging oncogene story and its implications for other programmes. "In a sense we anticipated these breakthroughs", Fischinger says, noting that approximately \$40 million was spent on this work in 1982 and \$43 million is to be spent in 1983.

But if NCI allows spending on oncogene research to expand naturally, does this mean less prominence for important traditional fields such as chemotherapy? DeVita says that some other work must obviously go, given the fact that NCI is unlikely to receive any budget increases in the next few years. He notes that chemotherapy has been cut by about 30 per cent in the past six years on scientific grounds: "some things we didn't need to do any more", he says. And, as explained by Alan Rabson, director of NCI's division of cancer biology and diagnosis, "if you understand oncogenes you may learn where to go in chemotherapy. It may open up whole new areas of chemotherapy".

But NCI's experience in another area of cancer research may be making DeVita wary of establishing a new programme or making other institutional changes. In 1964, NCI established what came to be called the Special Virus Cancer Program — a targeted research effort based on the belief that the discovery of a human cancer virus was just around the corner. In time, the programme got a bad name — it attracted money and policy attention but real scientific progress did not keep pace.

In due course NCI's Cancer Advisory Board appointed a group to review the Special Virus Cancer Program. Headed by Norton Zinder of Rockefeller University, it concluded that the programme had indeed gone too far. It has since been broken up and integrated into other work at NCI. Some of the work, however, contributed to the recent outbreak of research on human oncogenes since most, if not all, of them are closely related to viral oncogenes. What is more there is at last firm evidence of a human cancer caused by a retrovirus.

Today, Zinder comments on these newer, more profound breakthroughs. "In three years I can see there might be a need for a special program, but they're still exploring the basic implications." Using an analogy from American baseball, he said, "we're still running to first base with a good throw coming in from third". Clearly, NCI's system has more steps to take before anyone will feel confident that they can hit a home run and recommend a special, targeted programme.

Deborah Shapley

Spanish science policy

Socialists name science team

Barcelona

In Spain, the new socialist government is likely to spend more on science and technology than its predecessor, at least if its election promises are carried out. Since the elections on 28 October, the Spanish socialist party (PSOE) has an absolute majority in both houses of parliament for the next four years. Its leader, Felipe González, is the first socialist prime minister in Spain in peacetime.

In the slogans used by PSOE during the election campaign, the main proposal was a "change" but a moderate change. That was evident in the science and technology programme, which was uncontroversial enough to have been supported by any of the major parties contending the elections.



Felipe González, committed to doubling the share of the GNP to go to scientific research

An increase in the present level of spending in research and development is proposed in the socialist programme. An increase in public research funding should raise spending on research and development as a proportion of the gross national product (GNP) from its present 0.4 per cent to 0.8 per cent in the next four years, an average annual increase of 22 per cent. Besides these financial measures, the programme contains a number of more general proposals, including the definition of priorities, reform of the administrative structure of research centres and a reform of the situation of research staff.

The new Minister of Education and Science, Jose María Maravall, is a sociologist. From now on, there will be two people with direct responsibility for universities and research: Dr Carmina Virgili, a professor of geology, has been appointed Secretary of State for Universities while Dr Emilio Muñoz, a membrane biochemist, is the new Director General of Science Policy. A detailed programme for the next four years has been announced.

Pedro Puigdoménech

Nuclear physicists in Italy

Give us the money and we will deliver the goods

Rome

THE invitation ran: "*The Symposium will be held in the HONOUR ROOM of the PALACE of the KNIGHTS of the EQUESTRIAN ORDER of the HOLY SUPULCHRE of JERUSALEM near St Peter's Square*". A meeting of theologists? A canonization? Not at all. In fact, this was a pre-Christmas gathering of high-energy physicists — including Nobel laureates Abdus Salam, T.D. Lee, J.W. Cronin, Val Fitch and Sam Ting — all aiming to help their Italian colleagues lobby for reform of their principal institution, the National Institute for Nuclear Physics (INFN).

This seems to be the way things are done in Italy — plenty of pomp and circumstance — but it is emphasized by the particular style of INFN president, Professor Antonino Zichichi. Not content with the meeting itself (concerned with the international role of INFN) Zichichi also arranged an audience with the Pope and a meeting between four of the Nobellists and the Italian President, Sandro Pertini. But the highlight of the meeting appeared to be unplanned and unexpected. Claudio Villi, an ex-president of INFN and professor of theoretical physics at the University of Padua, slammed the government of his day (he was INFN president from 1970 to 1975) for "complete indifference to research". They could, he claimed, have found the heavy lepton (the tau, a heavy version of the electron and muon discovered at Stanford in 1976) and the J/psi (the first direct evidence for the existence of the charmed quark, discovered in 1974) well before the Americans, if the government had listened to their pleas. The discovery of these particles helped to change the face of physics in the 1970s, and it clearly would have been remarkable if Italy had first shown them to the world.

Villi may be accused of some Latin hyperbole, but there was clearly something in what he said. The tau and the J/psi were discovered using colliding beams of electrons and positrons. (Ting also discovered the J/psi in another way, but the method was almost a dead-end.) But the technique for colliding electrons and positrons was first developed in Italy, under the guidance of Austrian-born Bruno Touscek (now dead). The means was a ring called Adone, constructed outside Rome (at Frascati), which was taking data in 1970 — but at 3 GeV, an energy just below the threshold for J/psi. Zichichi was always pushing for higher energy, said Villi, and also for a big magnetic experiment to detect heavy leptons. (Zichichi in fact developed the technique for observing such particles used later at Stanford, using an acoplanar electron and muon of opposite sign.) If the

money had been available to stretch Adone, or early support had been given to Superadone (a rejected project for a 12 GeV machine), Italy might have taken a world lead . . .

There was strong support for Villi's view that physics should receive more support and Zichichi was naturally delighted. "This will give me power with the parliamentary commission". It was not just Villi: the meeting demonstrated the scale of international collaboration in physics. And this was the key problem facing INFN: that it suffered from bureaucratic civil-service controls and career structures which meant it was difficult to employ foreigners, even for short periods, and difficult for Italians to work abroad. But Adone was now lost to high-energy physics (devoted to synchrotron radiation) and it was essential that INFN look outward.

Other major difficulties with the INFN structure, Villi said last week, were that INFN careers were not recognized by the universities, so it was difficult to make transfers between say a professorship and INFN and that the pay for senior technicians and engineers is so low — relative to industry — that it was almost impossible to find qualified people to build experiments (mainly large particle detectors) to run on accelerators abroad. 700,000 Italian lira a month (£3,700 a year) is "indecent" says Zichichi. According to another respected Italian physicist, Ugo Amaldi, Italian teams designing detectors for LEP, the big electron-positron collider due for construction at the European Centre for Nuclear Physics (CERN) near Geneva were being severely hampered. French teams, for example, had ten times as many engineers as the Italians, and the Italian contribution to LEP was suffering. Zichichi also faces a danger: that the government will insist that the president of INFN be nominated by the government, rather than elected by the physicists as at present. This of course would reduce his power. But Zichichi is hopeful that the seminar and its fall-out will have influence. The present regulations are "fantastic, stupid" said Zichichi. He wants a structure in which INFN with its 1,000 staff and 1,000 unpaid collaborators can define its own regulations and rules, fix its own salaries (in agreement with the unions) and perhaps employ scientists on five-year renewable contracts to give flexibility and help push salaries up. He also wants an increase in the INFN budget from 75,000 million lira for 1983 to 120,000 million for 1984. "If we succeed in this we break the ice for the support of all basic science in Italy", he says. It would be extraordinary but it is just possible. **Robert Walgate**

Proprietary rights to cell lines

Roche pays up

Washington

THE legal battle over ownership of the KG-1 cell line was settled out of court last week by Hoffman-LaRoche and the University of California (Los Angeles). The university will receive an undisclosed "monetary consideration" for Roche's use of the cells, in conjunction with Genentech, to produce interferon. The cell line was isolated by two researchers at the University of California at Los Angeles, Dr David Golde and Dr Philip Koeffler.

Under the terms of the settlement, the parties are subject to a court order forbidding disclosure of the financial agreement. Charles Townsend, attorney for the university, said his clients are "very pleased indeed" with the settlement. As for its confidentiality, he said, "one might assume that any party that pays money would not want to have the amount revealed".

Although the monetary arrangement is between Roche and the university only, Golde and Koeffler, who were also named in the suit, are also expected to share in the award in accordance with the university's intellectual property policy. Townsend said that provides for a 50:50 split between the university and a faculty inventor, less a 15 per cent administrative fee for the university.

Golde refused to comment on whether he would share in the settlement, which he said was a "round figure" and "what I consider to be a large sum — though I don't know if anyone else would".

The agreement ends a three-year legal battle that became increasingly acrimonious, with each side countersuing and adding new accusations regularly. Most recently, Roche charged Golde with slander and defamation for allegedly calling Roche "crooks" who had "hijacked" his cell line; the university countered that Roche had "wrongfully interfered" with Golde's and Koeffler's professional relationships with the National Cancer Institute by spreading information "calculated to demean, misrepresent and disparage" them to the staff of the institute.

The seeds of the dispute were sown in 1978, when Golde supplied Dr Robert Gallo, a National Cancer Institute researcher, with the KG-1 cells. In a letter written to Gallo at the time, Golde said the cells "are made available exclusively to your laboratory for collaborative studies".

A year later, Gallo reported that the cells showed some promise for interferon production and, on 2 July 1979, Golde gave Gallo permission to investigate this.

At this point, the story become cloudy. Either directly or indirectly, Gallo then apparently made the cells available to Roche. Golde says his permission to transfer the cells to Roche was never sought, and that his understanding with Gallo that the

cells were not to be passed on remained in effect. Roche subsequently entered into an agreement with Genentech for the production of interferon from the cells.

The first legal action came in September 1980. Roche, in a pre-emptive strike, sued Golde, Koeffler and the University of California, seeking a declaratory judgement that the university had no claim on or interest in the cells or Roche's use of them. The university countersued, accusing Genentech of unlawfully using its property and Roche of unlawfully taking it.

The case was to go to trial in May. According to attorneys for both sides, there

was no one event that led to the settlement last week; Golde speculates that Roche decided to settle because of "worries about bad publicity of a trial".

The settlement makes no determination of facts with regard to ownership of the cells, but Golde believes it will nevertheless have "a profound effect on interactions between educational institutions and the private sector". The university successfully maintained its claim to property rights, Golde says, while maintaining as well the rights of researchers to share that property freely and openly for scientific purposes.

Stephen Budiansky

British universities

Tenure guarded

THE British government's decision to spend £100 million over the next three years on support for information technology and bringing new blood into universities received a cautious — not to say suspicious — welcome from the Association of University Teachers (AUT) at its recent council meeting. The caution stemmed from AUT's fear that many of the appointments to be made under the new scheme may be subject to waiver clauses denying security of tenure to those appointed, and thus undermine AUT's attempts to preserve that principle.

According to AUT, some 3,000 jobs in academic and related categories have been lost in Britain since the cutbacks in recurrent grants announced by the University Grants Committee in 1981: so far the redundancies have all been achieved through voluntary means, but recent proposals from the Committee of Vice-Chancellors and Principals indicate that attempts will be made to dismantle tenure in order to clear the way for further redundancies. The battle for security of tenure for full-time posts now seems likely to be fought at the University of Aston in Birmingham. But apart from this main battle, AUT is convinced that an estimated 10,000 contract researchers in universities and other institutions are being denied a reasonable degree of job security under existing arrangements and so are being deterred from pursuing a career in research. In particular, part-time researchers who are paid by the universities themselves seem to be getting a poor deal.

Under a new career structure agreement made 18 months ago, researchers funded by the Medical Research Council, on the other hand, seem to get rather a good deal. Contracts are either for a fixed period — non-renewable except under special circumstances — or for permanent tenured posts. AUT feels this arrangement avoids the situation in which quite senior researchers, accustomed to having their contracts renewed regularly, suddenly find themselves out of a job when funding dries up. But the Medical Research Council is unusual in this respect: most of the researchers funded by the Science and Engineering Research Council, for example, are on short-term appointments.

The Science and Engineering Research Council, in collaboration with the University Grants Committee, will be dispensing a large proportion of the money now available for information technology and for bringing new blood into the universities, apparently under existing rules. AUT now wonders if the existing rules are good enough, and is to pursue this point with Mr William Waldegrave, minister for higher education.

Tim Beardsley

Nuclear waste

Congress digs deeper

Washington

A SUSTAINED five-year effort to establish a national nuclear waste policy has finally born fruit. In the waning hours of the lame-duck session, Congress compromised on legislation that calls for the construction of two geological repositories, a "monitored retrievable storage" facility and at least one interim storage site for spent fuel as a stopgap. Construction of the first repository could begin as early as 1989 if all goes according to plan. (The bill was passed through a parliamentary manoeuvre that circumvented usual House procedures which had seemed certain to kill the bill by delaying action beyond the few hours left before adjournment.)

The most controversial provision in the compromise — and a significant victory for the nuclear power industry — is the federal government's acceptance of responsibility for spent fuel that exceeds the on-site storage capacity of reactors. The industry has maintained that unless the government provides this interim "away-from-reactor" (AFR) storage of spent fuel, some

of a repository will be subject to environmental impact review and to legal challenges. And, following the threat of a filibuster from Senator William Proxmire, whose home state of Wisconsin is considered to be a prime site for a granite repository, a very strong role for state governors was written into the site selection process: a vote of both houses of Congress will be required to override a governor's objection to the siting of a repository or AFR storage facility in his state.

The bill sets a tight schedule for the site selection and construction of the deep geological repositories, which are generally considered to offer the most permanent solution to nuclear waste disposal. The Department of Energy will have initially to pick five sites for possible study; by 1985, the list will be narrowed to three which will be characterized in detail. The President is required to make a final selection for the first repository site by 1987, with a decision from the Nuclear Regulatory Commission by 1989 on whether to issue a construction permit. A second repository site will be chosen three years later.

All operating and construction costs are to be covered by a fee on nuclear generated electricity, initially set at 0.1 per cent per kilowatt-hour. Military waste is not covered by this bill.

The bill also authorizes research and development on waste packaging, on repository design and on increasing on-site storage capacity for spent fuel at reactors. Included in the research and development programme is construction of a small test facility which will contain up to 100 canisters of high-level reprocessed waste or spent fuel and which will be used to study how well packaging and geological repositories can contain radionuclides. An unanswered question is whether the site selection process for the full-scale repositories may be on too fast a schedule, given the overlapping schedule for the research and development programme: the test facility is scheduled to begin operation by 1990, one year after construction of the first repository is to begin.

Stephen Budiansky



of the older commercial reactors will have to shut down within a few years. Environmental groups have opposed AFR, fearing that it will become the *de facto* permanent storage facility.

Environmentalists did win on two other points. First, any final decisions on siting

X-ray satellite

Plea for summer launch

EUROPE'S X-ray astronomers face an uphill struggle to ensure that EXOSAT, the only X-ray observatory in sight, is launched by next summer. A recent decision by the council of the European Space Agency (ESA) has dashed hopes of an immediate move to launch the satellite either on the next Ariane (L6) or on a Thor-Delta rocket provided by the National Aeronautics and Space Administration (NASA). ESA seems bent on waiting for the scheduled launch on the delayed Ariane L7.

EXOSAT was originally to have been aboard Ariane L6, which was to have been launched in September 1982. The failure of Ariane L5, combined with ESA's decision to launch a communications satellite on the delayed L6, meant that astronomers face an extra year's wait for their launch, with the possibility of degeneration of equipment and the loss of experienced personnel.

Last November, at the request of various research groups (see *Nature* 18 November 1982, p.207), ESA's Science Programme Committee asked the council to "take all actions necessary to secure a successful launch of EXOSAT before the closure of the summer launch window". This unspecific demand was intended to encourage the council to allow preliminary official contacts with NASA about the Thor-Delta option as well as to consider the launching of EXOSAT on Ariane L6 (now scheduled for late May). The committee's choice of words was also meant to minimize the political embarrassment caused by the possibility of a switch from a European to an American launcher.

At its meeting in December, however, ESA's council postponed the decision. As a result, Professor K.A. Pounds of the University of Leicester, a member of the Science Programme Committee, has sent a telex to European astronomers asking them to lobby council members before the next meeting on 19 January. And the Space Science Advisory Committee — a group of scientists that advises ESA — met just before Christmas and sent a resolution to ESA's Director General urging specifically that the Thor-Delta option be explored.

The timing is critical. NASA has a Thor-Delta immediately available for EXOSAT (which was originally designed for a Thor launch), but will require an official request by February for a summer launch to be possible. Ariane L6 is scheduled for launch in late May and, given a turn-round time of about eight weeks, Ariane L7 (and EXOSAT) could only just be launched by the end of the summer "window" in August. Astronomers fear that political pressures and the obligations on Ariane from a queue of commercial satellites could delay EXOSAT into 1984. **Philip Campbell**

UK student maintenance

Loan scheme takes shape

A NEW scheme that would partially replace the British system of maintenance grants for students in higher education with government loans is included in proposals which have been drawn up by Mr William Waldegrave, junior minister for higher education. Under the proposed scheme half of the cost of students' living expenses would be given as a grant, and half would be met by a loan to be repaid over a number of years when the recipient was in full-time employment. The controversial parental means test is, however, retained.

It has been clear for some time that the government favours some restructuring of the student support system, but earlier proposals drawn up by Mr Waldegrave's predecessor, Dr Rhodes Boyson, were shelved, ostensibly because of cost. The new proposals — which have not been finalized and which will have to obtain treasury and cabinet approval before seeing the light of day — will also entail an extra short-term cost, but may be politically more acceptable. It has been claimed that the cost of student maintenance, at present £600 million per year, could be halved under the new scheme.

Although it seems to be accepted that interest on the proposed loans would be paid at below commercial rates, many of the fine details of the plan, such as arrangements for repayment when ex-students are unemployed or are on low incomes, have yet to be decided. Other possibilities considered in the new proposals are suggestions that the grant/loan scheme should apply to students on a wider range of courses than are at present covered by maintenance grants, and that the age at which students qualify as "mature students" and so become independent of parental contributions should be reduced from 25 to 21.

The government is extremely unlikely to make public its detailed proposals on a contentious issue such as this in an election year. The most that can be expected is that if the nuts and bolts of the scheme can be sorted out and meet with political approval, some sort of commitment to a loans scheme will be included in the Conservative Party's election manifesto. The new scheme has, not surprisingly, been greeted with consternation by the National Union of Students. **Tim Beardsley**

United States in the Arctic

Congress warms to the task

Washington

THE United States came a step nearer a comprehensive policy for research in the Arctic when the lame-duck session of Congress, under pressure from the Alaska delegation in the Senate, at last passed an Arctic Research and Policy Act.

Scientists from other nations have long complained at the lack of central research efforts in the Arctic by the United States and Canada, the two largest Western nations with a stake there. Tore Gjelsevick, director of the Norwegian Polar Institute in Oslo, feels that the Soviet Union has the largest, most centralized and most technologically advanced effort in the Arctic. He estimates that there are 20,000 Soviet scientists working on Arctic problems and that the work includes both the area off the Soviet coast and the central Arctic Ocean region.

The United States government set up an inter-agency committee in 1968, and a second committee was set in hand by the National Security Council in 1971 but never actually met. But even the discovery of oil at Prudhoe Bay in 1968 and the plan to build the Alaskan pipeline failed to bring greater centralization or coordination to Arctic research. The new bill, however, would create an Arctic Policy Council with considerable independence. It would have five members: a "chairperson", two Alaskans and two Arctic research experts nominated by the heads of government

departments. The council would guide a new Arctic Research Commission which would be funded with \$25 million in federal funds and additional amounts from Alaskan state oil revenues.

The bill is a by-product of the fight by Senator Stevens and his colleagues to keep open a Naval Arctic Research Laboratory (NARL) which the Navy would like to close. The laboratory is sited at Point Barrow, and is a local employer of some importance. It remains politically popular in spite of its large size and somewhat unclear purpose.

The original version of the bill would have set up an Arctic Research Council consisting of the heads of cabinet departments whose job it would be to decide whether NARL should be kept open and to take in revenue from Alaskan oil development to provide the funds to keep it going. The new version of the bill, however, is far more concerned with Arctic research policy in general, and thus closer to the recommendations made by polar scientists through the US National Academy of Sciences' Polar Research Board.

The new commission will decide the fate of NARL and will have the job of running it if it is kept open. So US Arctic research seems to be gaining something after all, even from the bottom of the pork barrel.

Deborah Shapeley

Consequences of nuclear warfare

SIR — Nature has from time to time criticized individual scientists and groups of scientists for their activities in behalf of peace and disarmament, especially in respect to nuclear weapons. But even this back-ground left us unprepared for the tone and content of Vera Rich's article (28 October 1982, p. 769). Basing herself on an article by E. Teller published in Reader's Digest, she endeavours to represent nuclear war as somehow survivable and thus perhaps not so different from pre-Hiroshima warfare. She speaks of a "swift return to quasi-normality after nuclear bombardment". Referring to Teller's Reader's Digest article she writes that "[He] attempts to refute some of the more fantastic myths of what nuclear war would entail. [He quotes] instances from Hiroshima and Nagasaki: 'bridges open . . . a day after the blast, trains ran on the second day, streetcars . . . on the third'".

Her terminology "fantastic myths" betrays her own outlook. It betrays also a failure to understand some indisputable and obvious aspects of the situation. Now, unlike in the days of Hiroshima and Nagasaki, huge numbers of nuclear bombs could saturate a whole country. This is the age of the new word "overkill". Hiroshima and Nagasaki, hit by what would be very small bombs, could receive aid from elsewhere. Now there would be no elsewhere. The utter destruction which would ensue would permit no return to Rich's "quasi-normality".

Rather than taking as authoritative an article published by a single American physicist in Reader's Digest, we would rely upon the resolution adopted virtually unanimously by the (US) National Academy of Sciences (regrettably never published in Nature) on 27 April 1982. We cite only two of its findings: "science offers no prospect of effective defense against nuclear war and mutual destruction, . . . a general nuclear war could kill hundreds of millions and destroy civilization as we know it".

Unanimous endorsement for the entire resolution came in June 1982, from the Executive Committee of the American Physical Society and also from the Academy of Sciences of the Royal Society of Canada.

Eminent physicians have made it absolutely clear that the health care delivery systems would be destroyed by a nuclear bombardment. Death would be widespread. Those who survived the explosion would suffer horribly until their brief post-bomb existence ended.

This was agreed unanimously and documented scientifically by a panel led on the US side by Professor Lown of Harvard and on the Soviet side by Academician Chazov, Deputy Minister of Health of the

Soviet Union. The panel was televised nationally by the Soviet television network and in the United States by the Public Broadcasting System. Information concerning transcripts is available from the Public Broadcasting System, 609 Fifth Avenue, New York, NY 10017, USA.

ED BARBEAU
ERIC FAWCETT
TERRY GARDNER
BRYDON GOMBAY
LEE LORCH
DEREK MANCHESTER
DEREK PAUL
METTA SPENCER
LYNN TRAINOR

(Directors, Science for Peace)

Science for Peace,
Physics Department,
University of Toronto,
Toronto, Canada

VERA Rich's article was not based on Teller's Reader's Digest article but on a TASS complaint thereof — Editor Nature

Asteroids or comets

SIR — I appreciate Peter Smith's comments on the rigid anti-catastrophism in geological thinking (Nature 300, 217; 1982). However, he erred in attributing to Alvarez et al. the idea of a terminal Cretaceous cometary impact.

In fact, those authors postulated asteroid impacts, and argued strongly against the cometary idea, which was suggested by H. Urey (Nature 242, 32, 1973) and more recently by myself (Nature 285, 201, 1980). The two scenarios are fundamentally different; one postulated instant extinction (10^0 yr), the other less catastrophic (10^4 yr).

KEN HSU

Eidgenössische Technische Hochschule,
Zürich, Switzerland

Alternate culture

SIR — Both the authors (L. LeShan, H. Margenau) and the reviewer (M. Jahoda) of the book Einstein's Space and Van Gogh's Sky (Nature 25 November 1982, p. 385) clearly think that the word "alternate" has the same meaning as "alternative".

Since the battle to preserve the clear and valuable distinction between the meanings of these words has already been lost in the United States, it is probably too late to arrest, in this particular, the slide towards illiteracy in this country. However, your journal enjoys a wide circulation, so that my plea that scientists should occasionally consult their dictionaries may hopefully have some effect.

D.F. LAWDEN

Department of Mathematics,
University of Aston,
Birmingham, UK

Polywater reviewed

SIR — I have only now managed to obtain Felix Franks' book Polywater, published over a year ago (see Nature 5 November 1981)¹.

Even the picture on the cover reveals the intention of the author to show that the story of the investigation of polymeric water can be compared with a burst soap bubble. The book pays little attention to the scientific objective of the investigation proper, to its genesis, or to the results obtained.

In his book, Dr Franks appears arrogant enough to judge and teach, but much of what he writes proves incompatible with such a pretension. In describing the negative aspects of the subject the author exhaustively uses the popular press, including newspapers, while disregarding a number of serious scientific journals. Thus, on page 134 he writes: "Yeland Allen and Peter Kollman stated quite explicitly that the position they now adopted was: 'clearly opposite to that in our original work'. During the whole polywater controversy this was the only (sic) honest statement by a group of workers that their own new findings had forced them to change their mind. Their article in Nature of 22 October, 1971 [ref. 2] contains no equivocations, no ambiguities, and no question marks".

Further on, on page 190 (see also page 140), the author writes, in complete contradiction with reality: "The manner in which Derjaguin wrote off ten years of work and publicity in two sentences (sic) was not easy to accept for those who had followed him, and led to some bitter criticism. Derjaguin, however, was geographically inaccessible. . ."

Inaccessible? Foreign scientists each year visit my department at the Institute of Physical Chemistry of the USSR Academy of Sciences, and we have received as guests Langmuir, La Mer, De Boer, A. Zettlemoyer, B. Pethica (twice), Parfitt and many others.

Anyone consulting our publications on the matter³ should be convinced that the above accusations are untrue. A study of our papers in Doklady Akademii Nauk (including the English edition) and cited in Nature, covering the detailed analysis of the composition of anomalous columns of water, shows that we did not repeat work carried out by other authors at an earlier date. (Delays in obtaining the final results in our laboratory were connected with problems in gaining access to certain items of equipment such as an electron probe.)

Samples of polywater obtained by about ten different techniques have been subjected to analysis. In a number of cases, Si and O₂ were found to be the only admixtures, proving that those samples were products of leaching of quartz in the water condensate. This is the most reliable confirmation of the need to explain why

the dissolving power of water formed by condensation of vapour is many times greater than that of liquid water at the same temperature.

One possible cause was identified in the work of M. Prigogine and J. Fripiat⁴. Nonetheless, in his book, F. Franks completely disregards the essence of the ingenious and seriously substantiated explanation as presented by M. Prigogine and he expresses (on page 141) his astonishment at this explanation (first published in 1971⁴, and more fully detailed in ref. 5). Franks considers it wrong to substantiate this explanation by performing the experiment in conditions favouring the appearance of the anomalous condensate. I find criticism of this kind astonishing.

In order to support his analogy with a burst bubble, Dr Franks has apparently tried to suppress even a hint to the positive results of those investigations. However, in a volume edited by Franks himself, as well as in the Adamson's course of lectures on "Physical Chemistry of Interfaces" those positive results are distinctly formulated. Among such results is the discovery that violations of the capillary condensation equation, which have been detected not only by Fedyakin but even earlier by Patrick, Wildson, Shereshevskii, Chmutov and others, are actually caused by dissolved admixtures (and only partly by the adsorption layers close to the liquid meniscus, taken into account in my work).

For an author undertaking to present an objective view of the problem, and emphasizing in support his non-involvement in the investigation of polywater, his competence and goodwill (see page viii), it is unpardonable of him to write that the discovery of the phenomenon of the anomalous formation of columns is exclusively due to Fedyakin (see page 130), and that I appropriated Fedyakin's discovery (sic) (see page 184). In fact the previously cited paper by Clifford contains quite a different statement (see page 109): "Certainly in 1934 Wildson discovered that the properties of water in glass capillaries, and particularly its vapour pressure, depended on whether the water is introduced from the vapour or the liquid phase. However, the systematic study of the anomalous properties of water condensed in glass capillaries or quartz capillaries begins with the rediscovery of the phenomenon by Fedyakin in 1961–62".

I have continued those investigations together with Fedyakin and many other colleagues, continuously citing Fedyakin's earlier papers. See also page 49 of the book, for my version of the matter.

Franks seems sorry that "Derjaguin, however, was geographically inaccessible". In order to help the matter, Franks assumes the role of a detective. Unhindered by the "geographical remoteness", he writes (page 191): "The fate of Zhelezhny, another of Derjaguin's junior

colleagues, is a classic in the 'Catch 22' tradition. After taking part in the early experiments, he was worried about the possibilities of impurities in the anomalous water and, on his own initiative, submitted a sample to an East German spectroscopist (anonymous, B.D.) for analysis. The results indicated appreciable levels of impurities. . . Derjaguin did not consider these experimental findings worthy of note in his reports. However, Zhelezhny's name was omitted from all further papers describing research on anomalous water. . . Zhelezhny was held responsible for not having done his job properly. He disappeared from his post at the Physical Chemistry Institute in Moscow. . ."

This deserves to be called a slanderous speculation. The scenario outlined bears little relation to the actual events. I would like to add that while being employed at the Physical Chemistry Institute, in the Surface Phenomena Department directed by me, Zhelezhny — after having done his "criminal" has written and defended a PhD thesis and completed a considerable part of his DSci. dissertation. Indeed, he "disappeared" from our institute, but then he similarly disappeared from two others in Leningrad and Moscow.

A logical question now is whether Franks has the right to judge and teach ethical and scientific norms to other scientists, and where in reality lies the so-called "pathology" of science. And finally, I should like to summarize (see Clifford's paper⁶) the positive significance of the research in "polywater". Had this research not been carried out, and the role of dissolved substances not been revealed and thoroughly studied, then the Wildson-Shereshevskii-Chmutov-Fedyakin phenomenon would remain a mystery, and the phenomenon of an enhanced solubility, which is observed during the condensation of water vapour, would remain unknown (it could, of course, be surmised from the practice of purification of chemical vessels by vapour treatment).

The work on "polywater" also stimulated the theoretical investigation of the structure of water on the basis of quantum-chemical computations (such as the work of Allen et al.). In conclusion, I point out that the opinions presented on pages 188–189 of Franks' book (without reference to names), do not correspond to reality; he says that "For instance, the group at Birkbeck College, themselves among the first 'activists' in 1967, suggested several years later that Derjaguin's discovery had been taken seriously not so much because of its intrinsic merit but mostly because of Derjaguin's professional standing as a member of the elite. They then expressed the view that 'much of the published work since then is careless, inconclusive, and often misleading, if not wholly conjectural'."

In fact, in spite of my publications of 1962–67, an explosion of active interest in

research in polywater took place only in 1969, when those "misleading works" had been repeated by many scientists.

Thus, there is no reason to consider the naivety of the scientists (see also page 68) as the cause of that interest. It would be more correct to say that nature prepared a malicious surprise. Of course, it is easy in retrospect to recognize the errors made by those doing research in a new field, and then to slap on the label "pathological science" (see Chapter 10).

Such retrospective findings comprise also D. Eisenberg's assertion contained in his review of F. Franks' book in *Science*: "Astonishingly few of the polywater scientists considered the implications of thermodynamics, even though they lead to an immediate negative conclusion on polywater. Since polywater has a lower vapour pressure than ordinary water, with which it is in equilibrium through the vapour phase, it follows from the first and second laws of thermodynamics that the polymer is a stabler form".

This reasoning ignores that in our experiments we have measured not the pressure of the polywater vapour, but that of a solution of a very weakly volatile "anomalous component" in the normal water. A reduced vapour pressure of that solution is in agreement with Raoult's equation. As regards the anomalous component itself, we have supposed (and still believe) that the products of its evaporation are not the molecules of ordinary water; therefore, there is no reason to believe that I and other scientists have ignored the second law of thermodynamics.

In conclusion, I'd like to point out that it is always safer to insure oneself against errors by working within an accepted field, without incurring the risk involved in scientific "pioneering". In this connection, having put forward the concept on the structural peculiarities of the boundary layers of liquids tens and hundreds of molecules thick as early as the 1930s, I have continuously met with scepticism. Recognition of my work in this specific field has come only after more than 35 years. It was a paper by the prominent English scientist, W. Hardy⁸, that urged me to make those investigations in spite of the fact that the attempt at experimental substantiation, as undertaken by W. Hardy, proved to be grossly erroneous.

B. DERJAGUIN

Institute of Physical Chemistry of the
USSR Academy of Physical Sciences,
Moscow, USSR

1. Ubell, R. *Nature* 294, 28 (1981).
2. Derjaguin, B.V. & Churaev, N. *Nature phys. Sci* 232, 131 (1971).
3. Derjaguin, B. & Churaev, N.V. *Nature* 244, (1973); *J. Colloid Interface Sci.* (1974).
4. Prigogine, M. & Fripiat, J. *Chem. phys. Lett.* 12, 107 (1971).
5. Prigogine, M. thesis.
6. Clifford, p.120, Ch.5 in *Water* (ed. Frank, F.);
7. Eisenberg, D. *Science* 213, 1104 (1981).
8. Hardy, W. *Boundary State* (1931).

UN makes sense on radiation

Among the good works the United Nations has attempted in recent decades, its committee on the effects of atomic radiation deserves a special cheer.

MOST of the work undertaken by the United Nations is for one reason or another controversial, which in itself is a good reason for welcoming the latest report from the UN Scientific Committee on the effects of atomic radiation, now published. Since the committee was set up in 1955, at the height of world-wide anxiety about the effects of radioactive fallout from nuclear weapons tests, it has acquired a remarkable reputation for fair dealing and impartiality. By good management rather than by good luck, each successive report in the previous seven to be published has added to the committee's reputation. During the period for which the committee has been at work, its membership has naturally changed several times. So, too, has the problem — fallout is now less on people's minds than the possible exposure of the general public to radiation from the operation of civilian nuclear power plants, for example. The committee's eighth report, the bulkiest so far, is both a vivid pointer to the way in which the problem continues to change and a stimulating guide to the way in which scientific developments in widely different fields — nuclear engineering on the one hand and genetics on the other — can be welded into a sensible discussion of an interdisciplinary problem.

The committee's objective from the start has been to compile nothing less than a global inventory of people's exposure to radiation from all sources, natural and artificial. The general conclusion remains what it has always been — that natural sources of radiation, cosmic rays and the decay products of naturally occurring radioactive elements, remain the chief cause of people's exposure to radiation and thus, potentially, of damage. Part of the interest of the eighth report is that it clearly shows how even in the apparently well trodden field of the assessment of radiation exposure from natural sources, the conventional wisdom keeps changing. The most important development since the committee's previous report in 1977 is the recognition of the importance of airborne radioactive nuclides, principally radon and thoron, as sources of radiation exposure.

Exposure to cosmic rays, the effects of which vary both with altitude and with geomagnetic latitude, are calculated to yield an average dose of external radiation exposure of 0.30 millisieverts a year, not very different from but more accurate than previous estimates. The average external radiation dose from naturally occurring radioactive elements of the uranium-238

and thorium-232 series is, however, calculated as 0.23 millisieverts while the internal radiation dose provided by the volatile members of these radioactive series is estimated at an average of 1.15 millisieverts a year (with radon five times as important as thoron). The recognition that radon and thoron significantly contribute to the natural exposure of people to radiation is not of course new, but the numerical estimates now given are striking. It is no wonder that several public authorities have embarked on detailed studies of the exposure of their populations to radiation on this account. The committee's wry comment that exposure to airborne radioactivity may be increased by measures taken to conserve heat in houses, which have the effect of reducing ventilation, is unlikely to have much influence on the way that people behave if governments encourage them to do so. The practical consequence of this development is that it modifies the baseline, the natural exposure of people to radiation, against which the dangers of exposure from artificial sources must be assessed.

Fallout, the chief reason for the committee's existence, appears now to be less threatening only because of the way in which the nuclear powers have mostly given up the testing of nuclear weapons in the atmosphere. In reality, the average annual radiation dose from fallout is now declining below one per cent of that from natural sources (compared with the peak of 7 per cent in 1963). The consequences of medical irradiation, whatever they may be, are on the average at least ten times as important. Even so, the committee's new calculations have the virtue that they take account of a greater range of fallout nuclides (a total of 21) and are better informed than previously of the variation of fallout deposition from one part of the world to another.

The most novel part of the committee's latest report is its calculations of the consequences of civil nuclear power for the exposure of the general population to radiation. Hitherto, most estimates of the consequences of the development of a civil nuclear power industry have been guesswork, based on estimates by the designers of nuclear plants of the radiation doses to which workers would be exposed. This eighth report from the UN committee includes an impressive compilation of operating experience at reactors scattered around the world. Most reactors and all reprocessing plants release substantial

amounts of radioactivity in the form of gaseous and volatile isotopes to the atmosphere (argon-41 and iodine-129 for example), as a result of which the exposure of the general population is increased.

What stands out from the committee's new report is what has been intuitively clear all along — that at least in the short run, the exposure of the world's population as a whole to the release of radioactive materials in this form is so far, in aggregate, less important than the exposure of particular groups of people occupationally exposed to radiation. For the time being, however, it is clear that the committee is straining at a gnat. Its estimate of the total collective dose in 1979 to people occupationally exposed in the worldwide nuclear power industry — 2,000 man-sieverts — is identical with the estimate of the increased collective dose to people from travel in high-altitude jet aircraft. (It is worth noting that the committee's estimates of the doses accumulated by workers at reactors and reprocessing plants are roughly equal, that uranium mining does comparatively well and that waste storage has so far contributed negligible amounts of radiation dose.) What the future holds depends on two unknowable considerations — the pace at which the nuclear industry will grow, and the degree to which nuclear plant designers will be required to moderate the exposure of people to radiation.

The committee's next report will be well worth reading on this point, as will be its hoped-for account of the resolution of the doubts that have arisen in the past two years of the relative importance of gamma rays and neutrons in causing biological damage at Hiroshima and Nagasaki. The issue is whether the carcinogenic effects of the two kinds of radiation were inaccurately assessed in the early studies of the consequences of these two explosions. A changed estimate would principally affect the calculation of the likelihood that people exposed to different kinds of radiation would contract some kind of cancer. Wisely, the committee keeps its counsel on this important point. In doing so, it defines an issue to which radiobiologists everywhere should direct attention. And, in the process, it offers to other committees of the United Nations an example of restraint that should be more widely followed. □

Ionizing Radiation: Sources and Biological Effects. United Nations Scientific Committee on the Effects of Atomic Radiation 1982 Report to the General Assembly, with annexes.

Darwin centenary year

Galápagos evolution continues

from Peter T. Boag

WITH the Darwin centenary year now drawn to a close, it is already clear that it has produced a rich harvest of new information on the man and his work. Of the many papers, books and symposia produced in his honour, Darwin himself might have liked most to attend the last public event — a symposium on 'Evolution in the Galápagos', organized by R. Berry and hosted by the Linnean Society on 8 December in London¹.

Before Darwin, there was little appreciation of the biological uniqueness of islands, as A. Cain pointed out in his historical introduction. Cain ascribed this failure to the 'naive taxonomy' of the time. The lack of explanations for disjunct distributions of taxa or for the difference between convergence and affinity hindered 'amateur naturalists' such as Darwin when he set sail on the *Beagle* in 1831.

This theme was pursued further by F. Sulloway in discussing his own historical detective work on Darwin and the Galápagos^{2,3}. Sulloway dismissed as a myth the idea that Darwin had been converted to the theory of evolution by a 'Eureka!'-like experience in the Galápagos; instead his fuzzy grasp of systematics blinded him to the evolutionary significance of inter-island variation in the tortoises, mockingbirds and the now famous Darwin's finches. Darwin not only failed to separate correctly his finch specimens by island in the Galápagos, but also helped his shipmates eat their way through tortoise variants that were not collected again until years later.

What did happen was that in the second week of March in 1837, after his return to England, Darwin had a meeting with John Gould to discuss the *Beagle* bird specimens. Darwin was stunned by Gould's findings, and hastily scribbled the main points on the back of a scrap of museum paper. It seemed that, among other things, the motley group of birds Darwin had identified as a mixture of finches, wrens and blackbirds was in fact a closely related subfamily of passerines, new to science. Realizing the importance of knowing which forms had come from where, Darwin began the frustrating task of quizzing Captain Fitzroy and other shipmates, who had fully labelled specimens. Darwin was never entirely happy with the result, and perhaps for this reason omitted any mention of the finches from his *Origin of Species*. Sulloway suggested that this group of birds became closely identified with Darwin and evolutionary theory largely because of David Lack's book *Darwin's Finches*.

Recent geological work in the Galápagos and an improved understanding of the

plate tectonics of the region (B. Rosen and T. Simkin) show that the islands are younger than previously thought — no older than 3–5 million years. They have never been connected to the mainland, and remain among the 12 most active areas of vulcanism on Earth. The likely evolutionary age of several Galápagos vertebrates was reduced further still in J. Patton's discussion of genetic processes in Galápagos organisms. His electrophoretic studies of the finches, tortoises, lizards and rats suggest that with the possible exception of the land and marine iguanas, the small genetic distances among most taxa indicate evolutionary histories of under one million years^{4,5}. Biochemical characters produce phylogenetic trees very similar to those suggested by traditional taxonomic approaches but show a small and constant rate of change relative to changes in some morphological characters. This supports recent suggestions that evolutionary changes in structural genes occur in a clock-like fashion, while morphological changes can occur at different rates as selection acts on 'regulatory genes'.

Recent research on Galápagos land birds, and on Darwin's ground finches in particular (P. Grant), shows interesting developments in at least four areas, including the origin of Darwin's finches, the discovery of the first fossil finches, the study of contemporary evolution in the *Geospiza* and the confirmation of David Lack's model of adaptive radiation^{6,7}. A widely distributed South American emberizid, the blue-black grassquit (*Volatinia jacarina*), has recently been proposed by D. Steadman as the direct ancestor of all the Darwin's finches. Steadman has also discovered the first vertebrate fossil in the Galápagos, confirming the previous abundance of two finch populations rendered extinct since Darwin's time and suggesting that other finches as well as native rats went extinct before ever having been recorded^{8,9}.

Specific island populations of tortoise are being identified by T. Fritts, partly to aid captive breeding programmes by including animals of unknown origin located in zoos. The small genetic distances between tortoise populations have hindered this task, as has the apparently high degree of morphological convergence seen in different tortoise populations in similar environments. H. Snell has also looked at the similarity of various iguana populations in the Galápagos, as recent assaults by feral dogs have nearly obliterated two land iguana populations. Prompt work by Snell and his colleagues has led to the successful captive breeding

and reintroduction of these animals, in concert with programmes to eliminate the dogs¹⁰. Snell found that on the basis of morphology, land and marine iguanas are very similar to each other and different from mainland species. Patton had suggested that the large genetic distances indicated that they were unrelated and colonized the Galápagos separately a long time ago. But in a surprising announcement Snell revealed he had been studying what appears to be a land-marine iguana hybrid for three years. The hybrid shares features of each genus — it has, for instance, unwebbed front toes but webbed rear toes.

Y. Lubin and D. Porter discussed recent work on introduced invertebrates and plants. Lubin has been following the invasion of the fire ant (*Wasmannia*), which probably arrived in the islands in the 1930s, but is still dispersing at rates of up to 170 m yr⁻¹. As it enters an area, native ant species disappear, as do other arthropods such as scorpions and spiders. Porter pointed out that up to 26 per cent of the 738 known Galápagos plant taxa are now introduced, most within the past century¹¹. There is no sign that the rate of invasion or redistribution of introduced plants is slowing, and although many of the 'weeds' occur in disturbed habitats, several invaders threaten natural areas.

Finally, in a discussion of the role of man in the accelerated rate of 'evolutionary' change in the Galápagos, P. Kramer argued that in addition to trying to understand how native Galápagos plants and animals evolved and attempting to control introduced species, biologists should also monitor the process of invasion. Understanding why particular introduced plants or animals succeed or fail and how native species respond to introduced species may in the long run contribute as much to our understanding of evolution as research on native organisms¹². This brought home a bitter-sweet message about evolutionary research in the Galápagos: much less is known about evolution in the Galápagos than most people think, but Galápagos populations and communities are probably now changing faster than ever. □

Peter T. Boag is at the Edward Grey Institute, University of Oxford, Oxford OX1 3PS.

1. The symposium will be published in *Biol. J. Linn. Soc.* in 1983.
2. Sulloway, F.J. *J. hist. Biol.* 15, 1 (1982).
3. Sulloway, F.J. *Bull. Br. Mus. nat. Hist. (Zool.)* 43, 49 (1982).
4. Yang, S.Y. & Patton, J.L. *Auk* 98, 230 (1981).
5. Marlow, R.W. & Patton, J.L. *J. Zool. Lond.* 195, 413 (1981).
6. Boag, P.T. & Grant, P.R. *Science* 214, 82 (1981).
7. Grant, P.R. *Am. Sci.* 69, 653 (1981).
8. Steadman, D.W. *Proc. 8th int. Congr. Speleo.* 2, 549 (1981).
9. Steadman, D.W. *Trans. San Diego Soc. nat. Hist.* 19, 279 (1982).
10. Reynolds, R.P. *Not. de Galápagos* 36, 13 (1982).
11. Eliasson, U. *Not. de Galápagos* 36, 7 (1982).
12. Duffy, D.C. *Not. de Galápagos* 33, 21 (1981).

Early social evolution

Late Pleistocene life in Tasmania

from Geoff Bailey

In this issue of *Nature* (p.28), the discovery of a rich archaeological human occupation site at Fraser Cave on the Franklin River in south-west Tasmania is described. The site provides evidence of a specialized hunting economy, focused on the large wallaby, between 20 kyr and 16 kyr BP, and has an important bearing on a crucial period of human social evolution.

Populations of man with distinctively modern characteristics became widespread during the final readvance of the Last Glacial ice sheets (~30–20 kyr BP) and the period of their maximum extent (~20–16 kyr BP). After ~30 kyr BP anatomically modern populations of *Homo sapiens sapiens* with a technological expertise in stone flaking and grinding and firing of clay, and social and subsistence strategies informed by concepts of planning, storage and investment in future returns were widely in evidence. Personal decoration and artistic representation become increasingly prominent towards the Glacial Maximum, suggesting a fully developed capacity for symbolic thought and ritual expression and a greater awareness of individual and group status. Dramatic population expansion occurred, with occupation of Siberia and Australasia — and, at a later date when the continental ice sheets were beginning to retreat, northern Europe and the Americas — if not for the first time, then on a scale which dwarfs earlier evidence. There is greater diversity in the exploitation of food resources, especially from ~20 kyr BP onwards, including more intensive exploitation of foods from the sea and rivers, and new associations with plants and animals which foreshadow the development of full-scale agriculture.

Within the past decade our focus on this period has been progressively sharpened in three ways. First, a vast amount of new information about climatic and environmental change at the Glacial Maximum has become available. Extreme environmental variation was not confined to the vicinity of the ice sheets, but included global effects such as lowered sea level, reduced temperature and precipitation, extension of open tundra, steppe and grassland at the expense of forest, and expansion of arid desert margins, all of which would have had potentially large-scale disruptive effects on human behaviour.

Second, regional studies of archaeological settlement patterns in the Northern Hemisphere have built up an increasingly detailed and complex pattern of demographic and ecological

adjustments to these environmental effects^{1–3}. In some cases there was actual colonization of new territory, in other areas increased population density in already established zones of occupation, elsewhere decreased population density or complete depopulation. Specialization on a single prey species is found in many Upper Palaeolithic sites in Eurasia in response to the increased ecological specialization of extreme environments. The proliferation of art work may symbolize social adjustments to the new ecological conditions, for example the extension of reciprocal-exchange relationships integrating small groups scattered over large territories into social alliances as a form of insurance against the risks of economic specialization at low population densities^{4,5}. The apparent evidence of increased emphasis on marine resources, and at lower latitudes plant foods, may be interpreted as a broadening of diet in response to conditions of environmental or demographic stress.

Third, with archaeological discovery elsewhere, especially in Africa and Australasia, it has become increasingly apparent that Europe and western Asia were not unique centres of origin for demographic, economic and social change^{6–9}.

Hitherto Australia has seemed to differ from other centres by the lack of clear evidence for large-scale hunting — variously attributed to the coastal orientation of the earliest colonizers, unsuitable ecological conditions or poor archaeological preservation. The finds from the Fraser Cave, reported in this issue of *Nature*, significantly modify this picture. The geographical opportunism of the early Tasmanians in taking advantage of land connections exposed by lowered sea level and of hinterlands opened up by reduced forest cover, and their specialization on a single prey species, provide parallels with contemporaneous hunters at high latitude in the Northern Hemisphere. This pattern of convergence is reinforced by other evidence at lower latitudes, such as the similarities between parts of Australia, North Africa and the Near East in the probable shift to increased dependence on plant food during conditions of increased aridity after ~18 kyr BP, while the inclusion of sea and freshwater food resources in the diet occurred at an early period in Australia¹⁰.

These general comparisons underline the potential value of large-scale comparative research and the need for a programme of international collaboration aimed at giving

greater precision to patterns of local and regional environmental change during the onset of the Glacial Maximum, and the economic and social responses to them. This period represents a crucial threshold in human social evolution, during which the integration of characteristics fundamental to our subsequent success as a species first becomes apparent. It was also a period of extreme environmental variation, and more detailed comparative studies of human responses to these extremes would provide new insight into the constraints on human variability.

The Franklin River area offers an ideal regional laboratory for the analysis of long-term interactions between people and environment. There are many more caves than the Fraser Cave, and conditions of preservation are clearly excellent, with abundant stone industries and faunal assemblages. The area is a natural wilderness where the quality of the archaeological record has not yet suffered. However, the plans of the Tasmanian Government to flood the area as part of a hydroelectric scheme have caused international concern (see *Nature News* 300, 679; 23 December 1982). Scientific opinion was mobilized earlier in the year to meet this threat and a Senate Select Committee was set up to report to the Federal Government, which has ultimate responsibility for the area. The committee recently recommended that the caves of the Franklin River area were of such scientific importance that they should not be inundated. It also recommended that the Federal Government should fulfil its obligation under the UNESCO convention for the protection of the world's natural and cultural heritage and proceed with the nomination of south-west Tasmania for world heritage status. The Federal Government has not yet acted on these recommendations and if it does not do so, a unique opportunity to understand the early history of our species will have been lost. □

Geoff Bailey is Lecturer in Prehistory at the Department of Archaeology, University of Cambridge, Downing Street, Cambridge CB2 3DZ.

1. Marks, A.E. *Prehistory and Paleoenvironments in the Central Negev, Israel. Vol. II. The Avdat/Agave Area, Part 2 and the Har Harif* (Southern Methodist University, 1977).
2. Wendorf, F. & Schild, R. *Prehistory of the Eastern Sahara* (Academic, New York, 1980).
3. Bailey, G.N. (ed.) *Hunter-Gatherer Economy in Prehistory* (Cambridge University Press, 1983).
4. Gamble, C.S. in *Hunter-Gatherer Economy in Prehistory* (ed. Bailey, G.N.) 201 (Cambridge University Press, 1983).
5. Jochim, M.A. in *Hunter-Gatherer Economy in Prehistory*, (ed. Bailey, G.N.) 212 (Cambridge University Press, 1983).
6. Clark, J.D. *Man* 10, 175 (1975).
7. Mulvaney, D.J. *The Prehistory of Australia* (Penguin, Melbourne, 1975).
8. Klein, R.G. *Science* 172, 115 (1977).
9. Masters, P.M. & Flemming, N.C. (eds) *Quaternary Coastlines and Marine Archaeology* (Academic, London, 1983).
10. Bowdler, S. in *Sunda and Sahul: Prehistoric Studies in Southeast Melanesia and Australia* (eds Allen, J. et al.) 205 (Academic, London, 1977).

Reservoirs and earthquakes

Incident at the Aswan Dam

from R.D. Adams

LIKE the phenomenon itself, interest in reservoir-induced seismicity seems to run in cycles. In recent months the first detailed reports have appeared of a major new event at the Aswan Dam in Egypt, and there have been several re-appraisals of familiar events of the past. The effect of man-made reservoirs on local seismicity has been recognized since Carder's¹ early work at Lake Mead in Colorado, but the subject first attracted wide publicity in 1967 following the dramatic paper by Galanopoulos² giving details of the upsurge of seismicity following the filling of Lake Kremasta in Greece, when for the first time damage was attributed to an induced earthquake (see *News and Views* 295, 9; 1982).

It was soon realized that earlier large earthquakes at Xinfengjiang in China (1962) and Kariba in Central Africa (1963) were also reservoir-induced, and late in 1967 the largest induced event to date occurred at Koyna in India, with a magnitude of 6.5, causing more than 200 deaths and considerable damage.

Since then, more and more cases have been found, some spectacular and others evident only after careful statistical analysis. Recent compilations³ list more than 100 instances of reservoirs presumed to have induced earthquakes. Analyses of these cases suggested that minimum values of volume and depth were needed for induction, and that depth was by far the more important, reservoirs more than 100 m deep being much more prone to produce a seismic effect.

It was not therefore unexpected that Lake Nasser, behind the Aswan High Dam in Egypt, had induced no earthquakes since its initial impounding in 1965, for although it is the second most voluminous artificial reservoir in the world (exceeded only by Bratsk in the Soviet Union) its maximum depth is only about 72 m, and this was not reached until 1975. A recent article by Kebeasy and his colleagues⁴ shows, however, that an earthquake has now occurred — near the edge of the lake, 65 km to the south-west of the dam on 14 November 1981. It had a local magnitude of 5.6 (Helwan), body-wave magnitude about 5.25 (NEIS and ISC), was felt as far as Khartoum, 900 km to the south, and caused some damage at Aswan, although not to the dam or installations.

The event occurred at a considerable distance from the deepest part of the lake (65 km) and a considerable time after impounding (16 years since starting, 6 years since attaining full depth), and it is therefore natural to ask whether the earthquake was an induced event at all. Although there appear to be existing east-

west features in the epicentral area, and surface cracking was observed that could be associated with these, no instrumentally determined earthquakes have previously been located in the upper Nile Valley. Furthermore, Simpson's⁵ theory of mechanism for induced earthquakes suggests that the first events should be at some distance from the deepest part of the reservoir, and occur some time after impounding when the inducing effect of increased pore pressure first overcomes any inhibiting effect of loading. Thus a strong case can be made for an induced origin for this event, particularly as there appears to be a long-continuing aftershock sequence, which is another usual feature of reservoir-induced seismicity. These aftershocks were initially studied with portable recorders by the Geological Survey of Egypt and are now being monitored with a telemetered seismograph network in a joint venture between Helwan Institute of Astronomy and Geophysics and Lamont-Doherty Geological Observatory, New York.

Early studies of reservoir-induced earthquakes were mainly concerned with questions of seismicity, such as locations and numbers of events, and their relation to fluctuations in water level. The growing awareness of the role of fluids in earthquake mechanism, brought about by the advent of the dilatancy-diffusion hypothesis and experiments in seismogenesis by fluid injection, has recently rekindled interest in the subject.

It is therefore not surprising that as well as studies of new cases, many of the early induced sequences are being re-examined with newer techniques in the light of modern geophysical concepts, and in particular, detailed studies are being made of variations in position and mechanism with time.

The Kremasta sequence is the subject of a recent review⁶ which shows that the mechanism of the aftershocks changed after 9 months from normal faulting (which is the favoured mechanism for induction) to thrusting with a different orientation, and the suggestion is made that this change reflects the end of the induced phase and the resumption of normal regional seismic activity.

Another article⁷ continues the work of Simpson and his Soviet co-workers on activity near the Nurek Reservoir in Central Asia, which in contrast with Lake Nasser is high (315 m) with a small area and in a region of predominantly compressive tectonics. These induced earthquakes have not been particularly large, but have been carefully monitored and related to seasonal changes in water level. Following a rapid

rise in level in 1976 there was a spectacular increase in activity and detailed studies show that the foci do not correlate with planes associated with surface faults, but form short segments of alternating strike-slip and thrust faulting on a scale of a few kilometres.

A major aftershock of the Xinfengjiang event is studied by a group from the Seismological Bureau of Yunnan and Guangdong Provinces⁸ who use geodetic methods to obtain focal parameters, including seismic moment and stress drop, which complement well the conventional seismological investigations.

In yet another recent report⁹ the sequence following the Oroville, California, earthquake of 1975 is looked at in more detail, and it is shown that local seismicity decreases as the lake fills during winter and spring, and that the strongest shocks occur as the level drops during summer and autumn. The magnitude of these events seems to be decreasing with each seasonal drawdown. This was another case of late induction, with the main earthquake (magnitude 5.7) occurring nearly eight years after impounding.

As time goes on, more instances of slow-acting induction may be found, and more reservoirs such as Lake Nasser, assumed to be seismically inactive, may produce earthquakes. Continuing detailed study of such events should confirm whether this activity is or is not truly induced by the reservoir. It is generally accepted that induced seismicity only releases strain already stored in the region, perhaps bringing forward in time earthquakes which would have occurred in the future. In no cases should reservoirs increase the long-term seismic energy release, and it may be that an episode of induced seismicity will be followed by a compensating period of quiescence. Continuing monitoring of all existing reservoirs that have already induced earthquakes, and new ones that are likely to, should continue to produce scientific rewards as well as providing the information necessary to protect dams and associated engineering structures. □

R.D. Adams is at the International Seismological Centre, Newbury, Berkshire RG13 1LX and the Department of Geology, University of Reading, Reading.

1. Carder, D.S. *Bull. seism. Soc. Am.* **35**, 175 (1945).
2. Galanopoulos, A.G. *Ann. géol. Pays hell.* **18**, 281 (1967).
3. Perman, R.C., Packer, D.R., Coppersmith, K.J. & Knuepfer, P.L. *Rep. for USGS 14-08-0001-19132* (Woodward-Clyde Consultants, San Francisco, 1981).
4. Kebeasy, R.M., Maamoun, M. & Ibrahim, E.M. *Bull. int. Inst. Seis. Earthq. Engng* **19**, 155 (1981).
5. Simpson, D.W. *Engng Geol.* **10**, 123 (1976).
6. Stein, S., Wiens, D.A. & Fujita, K. *Earth planet. Sci. Lett.* **59**, 49 (1982).
7. Keith, C.M., Simpson, D.W. & Soboleva, O.V. *J. geophys. Res.* **87**, 4609 (1982).
8. Zhu, C., Liu, Y., Wang, C., Lu, R. & Chen, J. *Bull. seism. Soc. Am.* **72**, 1085 (1982).
9. Toppozada, T.R. & Morrison, P.W. *Calif. Geol.* **35**, 115 (1982).

Planetary exploration

Probing the lunar interior

from L.J. Srnka and S.K. Runcorn

WHETHER there is an iron core in the Moon, and, if so, what radius it has and when it was formed, has become a fundamental issue in lunar science. Up to the time of the Apollo 11 landing in 1969 it was assumed, for a variety of reasons, that the Moon had no iron core. The Moon's mean density (3.34 g cm^{-3}) is very close to that of olivine, it possesses no dipole magnetic field and its interior appeared, at that time, to have remained cold since its accretion. Although a suggestion¹ that the Moon has an iron core of radius about 350 km had been made years earlier, the argument was indirect, based on an explanation of the non-hydrostatic second harmonic of the figure by solid-state convection rather than by a primeval distortion based on finite strength.

The discovery of lunar palaeomagnetism in the lava samples returned from the Apollo 11 landing once again raised the issue in acute form. It seemed then that a primeval magnetic field generated in a core by the dynamo process was a reasonable explanation of the crustal remanent magnetization. The existence of a fluid iron core capable of generating a field probably 4.2 Gyr ago raises the fundamental question of what heat sources, following the Moon's accretion, could have melted the core to the centre and not, as previously supposed, only to about 300 km depth.

Direct evidence for the existence of a core is sparse — one seismic event and a value for the moment of inertia factor consistent with, but not requiring, a core.

Evidence for the existence of a core has been sought in magnetic-field variations but, because of the complexity of electromagnetic induction, the data obtained during the Apollo project continue to pose complicated physical questions that prevent full analysis. But impressive progress has been made, as reported in recent articles²⁻⁴ from a team led by C. Sonett at the Lunar and Planetary Laboratory, University of Arizona. Their re-analyses of ultra-low-frequency data (10^{-5} – 10^{-3} Hz) indicate that a metallic core (conductivity $\geq 10^2 \text{ S m}^{-1}$) cannot exceed 360 km in radius. Indeed, they argue that the data are consistent with a core-free Moon, because at these frequencies conductivity rises from insulating values at the surface to values which need be no more than 10^{-1} S m^{-1} at the centre. Such a conductivity profile can be explained by temperature rise and compositional changes over electrical semiconductors in silicate mineral assemblages. The conclusions disagree with those reached by C. Russell and co-workers⁵ at UCLA and the Jet Propulsion Laboratory in

California, who use different data and analyses and argue that a metallic ($\geq 10 \text{ S m}^{-1}$) core of at least $435 \pm 15 \text{ km}$ radius exists in the Moon.

Lunar induction data are measurements of the response of the Moon to external magnetic fluctuations, analogous to the geomagnetic depth-sounding technique first developed by Schuster⁶ in 1889 as a means to study the Earth's electrical structure. These fluctuations are caused by the flow of magnetized plasma from the Sun — the solar wind. The Moon encounters three flow environments in its orbit: the free-stream wind, the Earth's magnetosheath and the Earth's complex magnetotail (which consists of upper and lower lobes of opposite field direction, separated by the plasma sheet and neutral sheet).

Sonett and co-workers have used data from the Apollo 12 surface magnetometer and the Explorer 35 satellite in high lunar orbit, during periods when the Moon was in the supersonic solar wind or magnetosheath. Their technique resembles an electromagnetic TE mode⁷ scattering experiment, in which orthogonal frequency-dependent transfer functions are measured at the surface of the scatterer. The transfer functions are ratios of the magnitudes of the components of the surface field to those in the incident field (convected by the solar wind). If the incident wavelengths are sufficiently long (frequencies less than $1\text{--}3 \times 10^{-3}$ Hz for typical flow velocities), the scattering is in the Rayleigh regime and the dipolar response of the Moon is excited primarily. Unlike vacuum scattering, the induced field is here confined to the lunar interior by the hydromagnetic pressure of the solar wind and, fortunately, in this limit the asymmetry in the problem due to the lunar plasma wake can be neglected. Sonett *et al.* estimate the conductivity profile by forward computer modelling based on a 10-layer Moon; the preferred profile is one which matches the measured data to some degree of confidence. The maximum core (radius 360 km, conductivity $\geq 10^2 \text{ S m}^{-1}$) just fits with the 1σ confidence limit at 10^{-5} Hz, the lowest frequency studied and the one most likely to show a core signature (deepest penetration into the Moon).

It would be desirable to test this conclusion using the growing power of inverse theory. Hobbs⁸ used a linearized Backus–Gilbert technique on higher-frequency (5×10^{-4} – 4×10^{-2} Hz) lunar induction data, which places the scattering in the Mie regime. Hobbs found that such data give no information (that is, zero 'resolution') on the profile below 600 km. But the recent work of Hood *et al.*³, which

extends the useable data to 10^{-5} Hz and reduces the overall errors, is claimed to provide constraints on the profile to at least 1,350 km depth. The important question of true resolution in these data, perhaps in the Backus–Gilbert sense, should be addressed using inverse theory. This is work for the future.

The methods of Sonett and co-workers assume that the magnetometers on Apollo 12 and Explorer 35 were intercalibrated to better than one per cent, and that the solar wind impinges directly on the surface at Apollo 12. Examination has revealed that the magnetometers were not always in calibration, and also that the local remanent magnetic field at Apollo 12 site on Oceanus Procellarum could have deflected some of the solar wind and altered the spectrum of the incoming waves. Hood *et al.*³ claim to have accounted for these effects.

But an alternative induction experiment can be done that avoids these problems entirely: a single magnetometer in low ($\sim 100 \text{ km}$) lunar orbit can be used to measure the whole-body response of the Moon to the reversal in external magnetic field as the Moon passes from one lobe to another in the Earth's magnetotail. Russell *et al.*⁵ used data from the Apollo 15 and 16 subsatellites in this way. They found that the induced magnetic dipole moment is negative after such a passage, indicative of a diamagnetic response (exclusion of impressed fields) from the region within the orbits of the subsatellites. First interpretations attributed the diamagnetism to a lunar ionosphere lying under the subsatellites, but such an ionosphere has been shown to be unstable. Russell and co-workers now conclude that the response is due to a metallic lunar core, but admit that the limited data do not place constraints on its conductivity. If the Moon is magnetically nonpermeable (for example, no free iron in regions cooler than the Curie point), a core of radius $435 \pm 15 \text{ km}$ with conductivity $\geq 10 \text{ S m}^{-1}$ within an otherwise insulating Moon is claimed to be consistent with their data. If the Moon is paramagnetic, a larger conducting core is admissible since paramagnetism would mask its diamagnetic effect.

But the approach of Russell and co-workers has its own weaknesses: various effects due to the subsonic but often nonthermal plasma present in the magnetotail have not been taken into account; the very low-frequency (10^{-5} Hz) measurements needed to assure sufficient penetration of the reversed tail field into the conducting lunar interior cannot be obtained, since the Moon leaves the magnetotail too quickly; and the subsatellite data are equally explained by a larger central region with moderate conductivity (10^{-1} – 10^0 S m^{-1}) such as is expected for the lower lunar mantle.

The important questions concerning the existence and electrical properties of a lunar core could be answered (in principle)

using electromagnetic induction if high-quality data can be obtained at frequencies down to a few times 10^{-6} Hz. The short time overlap in the operations of the Apollo 12 and Explorer 35 magnetometers has led Herbert⁹ to suggest that data of this kind might be extracted from the extensive simultaneous Apollo 15 and 16 surface magnetometer measurements. Herbert has also sketched out a mathematical theory that describes how these surface measurements alone could be used to study the deep lunar interior. However, the TE mode transfer functions at these sites are known to have asymmetries in their horizontal components. These asymmetries are attributed to local subsurface electrical inhomogeneities, whose effects should be removed from the data before the new method could be implemented. It is hoped that the very long-wavelength data (frequencies below 10^{-5} Hz) would not be affected strongly by these local phenomena.

Planetary exploration has sometimes been criticized as not being 'fundamental' science. Perhaps the wealth of beautiful photographs obtained has distracted attention from the problems of fundamental physics raised by the data — if that is indeed the case the papers described here illustrate very well the sophistication of lunar physics. From the above discussion, it is clear that the issues will be resolved best if additional data are acquired, particularly in the low-frequency range. The project now being considered by ESA of a satellite in nearly polar lunar orbit at 100 km altitude, plus a relay satellite in lunar orbit at several thousand km altitude, would include sensitive three-component magnetometers as well as other sensors. The mission would have a duration of at least one year, so magnetic variations would be recorded while the Moon is in its various plasma environments. This would produce a new impetus to understand the physics of the interaction, and could shed light on the issue of a lunar core. ||

L. J. Srnka is in the Long Range Research Division, Exxon Production Research Company, Houston, Texas 77001 and S.K. Runcorn in the School of Physics, The University, Newcastle upon Tyne NE1 7RU.

Sexual systems

When to be which sex

from Robert M. May

NEW and more rigorous analyses of the evolution of sex and sexual behaviour have been the cause of much excitement in recent years. Two areas have attracted particular interest. One large body of work has sought to evaluate the circumstances under which sexual reproduction is likely to be more advantageous than asexual clonal reproduction¹⁻³. Another follows Darwin's lead in exploring the evolutionary basis of the diversity of mating systems and sexual behaviour found in vertebrate and other animal societies⁴. A third area of research has perhaps not received the attention it deserves — the consideration of the various forms that sexual reproduction may take.

As summarized by Charnov in his new book *The Theory of Sex Allocation*⁵, three main forms of sexual reproduction may be distinguished. Most familiar to us is dioecy, in which males and females are distinct and separate throughout their lives. Sequential hermaphrodites function as one sex early in life, and then switch to the other sex for the remaining part of their lives: the sequence where an individual begins life as a male and later becomes a female is called protandry and the converse sequence protogyny. In simultaneous hermaphroditism (often just called hermaphroditism), an individual produces both male and female gametes throughout its breeding lifetime.

This range of reproductive modes poses many evolutionary questions⁵. What is the equilibrium sex ratio — the fraction of offspring that are males — in a dioecious species? Which sex should come first, and when should the change occur, in a sequential hermaphrodite? What fraction of reproductive effort should be allocated to female versus male function in a simultaneous hermaphrodite? And, over all, what biological and environmental circumstances are likely to favour one particular reproductive mode over the other two?

The key to answering these kinds of questions was provided by Fisher⁶, who observed that, in general, natural selection will favour those females who apportion their reproductive resources equally between sons and daughters. If there were an 'unbalanced' population in which, for example, more resources were allocated to producing daughters than sons, then an individual female who changed to favouring production of sons would, due to the relative rarity of males, have a better than average genetic input into subsequent generations. Hence the propensity for the evolutionarily stable state to be one of 'equal resources'. Ghiselin², Hamilton⁷, Shaw and Mohler⁸, Leigh⁹ and others have

refined and elaborated Fisher's basic work, recasting his proposition in various ways. One particularly useful reformulation is by Charnov¹⁰, who begins by defining m to be an individual's average gain in fitness through male function (production of sons, pollen, or time spent as a male) and f the corresponding gain in fitness through female function (daughters, seeds, time spent as a female). The evolutionarily stable allocation of resources is then the one maximizing the product mf .

For dioecious organisms, the mathematical derivation of Fisher's result from Charnov's formula is simple but instructive. Let x be the fraction of reproductive resources allocated to production of sons, and $(1-x)$ the fraction to daughters. Then m is proportional to x and f to $(1-x)$. Although a variety of other constants are involved in the explicit formula for the product mf , the essential quantity to be maximized is just $x(1-x)$ as a function of x , and this quantity is at a maximum when $x = 1/2$. For species such as our own, females and males are sufficiently similar that equal allocation of reproductive resources translates into approximately equal production of male and female individuals. (The fact that at Princeton University, with a sex-blind admission policy, 36 per cent of the applicant pool and 37 per cent of the undergraduates are female says unfortunate things about social, as opposed to evolutionary, factors in allocation of resources.) As reviewed by Charnov⁵, there are many kinds of complications in dioecious species that can result in the need to maximize the product mf leading to very skewed values of the ratio between the number of males and of females. As one among many possible examples of such complications, Van den Assem¹¹ and Charnov *et al.*¹² have shown that parasitic wasps of the species *Lariophagus distinguendus* and *Heterospilis prosopoidis* tend to produce males from relatively small hosts and females from relatively large ones. As larger wasps emerge from larger hosts, the observed egg-laying behaviour would be expected if female wasps gain relatively more reproductive success from large body size than do males. Marshalling data that are directed towards testing this hypothesis, Charnov⁵ concludes they are suggestive but not conclusive. There are many other parasitic hymenopterans that attack a range of sizes of hosts, and where the sex ratio changes with host size. Quite apart from its relevance to the theory of sex allocation, such behaviour can — as recently emphasized by Waage and Hassell¹³ — introduce density-dependent

1. Runcorn, S.K. *Nature* **195**, 1150 (1962).

2. Hood, L.L. & Sonett, C.P. *Geophys. Res. Lett.* **9**, 37 (1982).

3. Hood, L.L., Herbert, F. & Conett, C.P. *J. geophys. Res.* **87**, 5311 (1982).

4. Sonett, C.P. *Rev. Geophys. Space Phys.* **20**, 411 (1982).

5. Russell, C.T., Coleman, P.J. Jr & Goldstein, B.E. *Proc. Lunar planet. Sci.* **12B**, 831 (1981).

6. Schuster, A. *Phil. Trans. R. Soc. A* **80**, 467 (1889).

7. Stratton, J.E. *Electromagnetic Theory* (McGraw-Hill, New York, 1941).

8. Hobbs, B.A. *Geophys. J. R. astr. Soc.* **51**, 727 (1977).

9. Herbert, F. *Moon and Planets* **23**, 127 (1980).

effects that significantly influence the population dynamics of the parasite.

Sequential hermaphroditism is widespread among invertebrates and fishes, and is known for a few plants. Ghiselin² and others have observed that natural selection is likely to favour such sex reversal over dioecy when an individual's reproductive success, or survival probability, or growth rate is closely related to age or size, and where this relationship is significantly different for males and females. Sequential hermaphroditism provides some exciting opportunities to confront evolutionary theory with observations and experiments.

In particular, many protandrous sequential hermaphroditic species exhibit two alternative life history pathways: a proportion, P , of individuals mature as females and remain so (called 'early-maturing females'); the remaining proportion, $1-P$, mature as males and then, at some well defined later time, reverse sex to become females for the rest of their lives. Some protogynous species exhibit a similar dichotomy, with the words 'male' and 'female' interchanged in the preceding sentence. In both cases, the evolutionarily stable value of P can be predicted on theoretical grounds, provided only that we know W , the fitness of a female (male) hermaphrodite relative to that of an early-maturing female (male). Charnov¹⁴ observes that for a protandrous hermaphrodite, the fitness associated with male function, m , is proportional to $1-P$ (along with other factors, assumed independent of P), and similarly, the fitness associated with female function, f , is proportional to $P + W(1-P)$. Thus the product to be maximized is $(1-P)[W(1-P) + P]$, leading to the prediction that

$$P = \frac{1}{2} [1 - W/(1-W)]$$

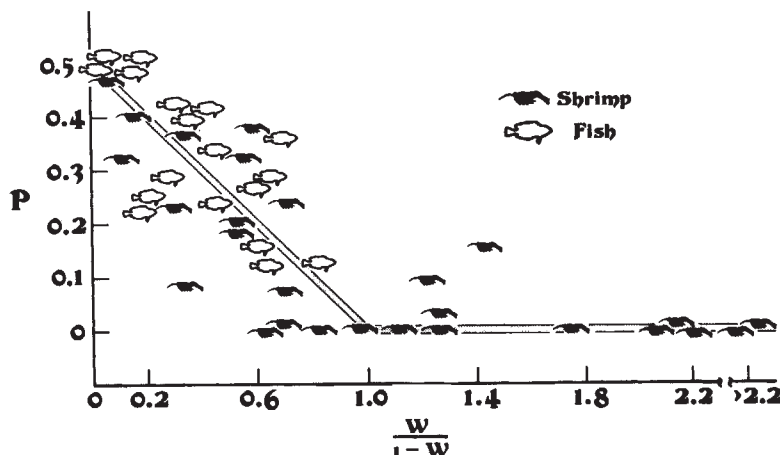
if $W < \frac{1}{2}$, and

$$P = 0$$

if $W > \frac{1}{2}$. The same expression holds, *mutatis mutandis*, for a protogynous species. Charnov's theoretically predicted relation between P and W is shown in the figure.

Many species of shrimp in the family Pandalidae are of commercial importance, in waters ranging from temperate to sub-Arctic. Although sex reversal is rare among crustaceans, most of these commercial species of shrimp are protandrous hermaphrodites. A good deal of information about them is available, by virtue of the fisheries. For some 27 populations of pandalid shrimp, Charnov^{5,14} has determined both the average proportion that are early-maturing females, P , and the fitness ratio, W (obtained from estimates of the average lifetime egg production of a hermaphrodite as a ratio to the lifetime egg production of an early-maturing female). These empirical values of P and W are plotted, as little black shrimp, in the figure.

In contrast with the protandrous



For protandrous shrimp populations and protogynous coral fish populations, the proportion P that do not change sex is displayed as a function of the appropriate fitness ratio $W/(1-W)$, as discussed in the text. The solid lines show the theoretical prediction. (From ref. 5, © 1982 by Princeton University Press. Reprinted by permission.)

shrimps, many parrotfish, wrasses and other fishes of tropical reefs are protogynous, changing their sex from female to male. These tropical fish species are all territorial, and it seems likely that this form of sex allocation arises because the larger size associated with older age classes confers a greater advantage on males (for defending territories) than on females. The detailed studies by Warner and Hoffman¹⁵ and others of the wrasses *Thalassoma bifasciatum* in the Caribbean and *T. lucasanum* in the tropical Eastern Pacific have shown the populations to comprise a mixture of pure males ('early-maturing males') and protogynous hermaphrodites — a mirror image of the pandalid shrimp populations. For these wrasse populations, estimates of the fitness ratio, W , between male hermaphrodites and early-maturing males can be made, albeit by somewhat indirect means^{5,15}. The resulting values of P and W for populations of *T. bifasciatum* on 18 different coral reefs of varying sizes are shown, as little open fish, in the figure.

I think the figure is rather beautiful. It indicates an encouraging agreement between the basic evolutionary theory and two very different sets of data.

In the third main form of sex allocation, simultaneously hermaphroditic individuals each allocate reproductive resources both to female and to male function. Charnov⁵ shows that such individuals will have higher fitness than dioecious ones when "male reproductive success shows a law of diminishing returns with the shunting of resources from female to male function" and when female reproductive success behaves similarly. If one, but not both, of male and female reproductive success exhibits diminishing returns, the outcome between hermaphroditism and dioecy will depend on the relative degrees of nonlinearity. A biological example to this effect has been provided by Heath¹⁶, who has suggested that the documented correlation between hermaphroditism and

brooding of offspring may arise because egg production is limited by the space available in the brood-pouch, so that excess resources are available for sperm production. Charnov⁵ shows how this qualitative idea can be elaborated, and indicates some possible quantitative tests. Things get yet more complicated when, as happens in many plants, the hermaphroditic individual fertilizes itself, resulting in inbreeding depression in the selfed offspring. Charlesworth and Charlesworth¹⁷ have made a start on the theory here.

Many people may be bemused by recent criticisms that evolutionary biology is largely a collection of 'Just So Stories', part of a capitalist-inspired 'adaptationist programme'. Charnov's monograph⁵ on *The Theory of Sex Allocation* testifies against this view, showing with many explicit examples how 'selection thinking' can lead to predictions that may be tested against experimental facts.

Robert M. May is Class of 1877 Professor of Zoology at Princeton University, Princeton, New Jersey 08544.

1. Williams, G.C. *Sex and Evolution* (Princeton University Press, 1975).
2. Ghiselin, M.T. *The Economy of Nature and the Evolution of Sex* (University of California Press, Berkeley, 1974).
3. Maynard Smith, J. *The Evolution of Sex* (Cambridge University Press, 1978).
4. Darwin, C. *The Descent of Man, and Selection in Relation to Sex* Reprint of 1st edn with an introduction by J.T. Bonner and R.M. May (Princeton University Press, 1981).
5. Charnov, E.L. *The Theory of Sex Allocation* (Princeton University Press, 1982).
6. Fisher, R.A. *The Genetical Theory of Natural Selection* (Oxford University Press, 1930).
7. Hamilton, W.D. *Science* **156**, 477 (1967).
8. Shaw, R.F. & Mohler, J.D. *Am. Nat.* **87**, 337 (1953).
9. Leigh, E.G. *Am. Nat.* **104**, 205 (1970).
10. Charnov, E.L. *Am. Nat.* **113**, 465 (1979).
11. Van den Assem, J. *Neth. J. Zool.* **21**, 373 (1971).
12. Charnov, E.L., Hartogh, R.L., Jones, W.T. & Van den Assem, J. *Nature* **289**, 27 (1981).
13. Waage, J.K. & Hassell, M.P. *Parasitology* **84**, 241 (1982).
14. Charnov, E.L. *Am. Nat.* **113**, 715 (1979).
15. Warner, R.R. & Hoffman, S.G. *Evolution* **34**, 508 (1980).
16. Heath, D.J. *J. theor. Biol.* **81**, 151 (1979).
17. Charlesworth, B. & Charlesworth, D. *Biol. J. Linn. Soc.* **14**, 57 (1981).

Pope Gregory's calendar

Keeping up with time

from David W. Hughes

WHEN astronomers use the term 'year' without any qualifying adjective they are referring to the tropical year, Y , the time between successive passages of the Sun through the vernal equinox. Unfortunately $Y = 365.24219879 - 0.000013T$ mean solar days, where T is the number of centuries that have passed since 1900.00. I say 'unfortunately' because it is obvious that a calendar year, for all practical purposes, must contain an integral number of days. So calendar years do not agree with astronomical years and to stop dates getting further and further out of step from year to year the number of days in a calendar year must be occasionally varied.

Julius Caesar introduced the so-called julian calendar in 45 BC. Every year that was exactly divisible by four was regarded as a leap year and thus contained 366 as opposed to 365 days. Unfortunately, the tropical year is not 365.25 days long but is 11 minutes and 14 seconds shorter; and this difference gradually mounts up.

By 1582 the julian system was about 10 days out. Easter in the Christian church is celebrated on the first Sunday after the first full Moon that occurs on or after the vernal equinox. This equinox was assumed to be 21 March. By 1582 the equinox was actually occurring on 11 March while Easter was still being reckoned from 21 March. If the julian calendar had been adhered to, Easter would eventually have been celebrated in the summer.

In 1576 Pope Gregory assembled a commission of astronomers, mathematicians and clergy to study methods for correcting the calendar. They eventually adopted a plan put forward by a Calabrian physician Luigi Lilio. The work of Lilio is discussed in a recent paper by Gordon Moyer in *Sky and Telescope* (64, 418; 1982). On 24 February 1582 Gregory XIII issued the papal bull *Inter Gravissimas*. Thursday, 4 October 1582 was to be followed by Friday, 15 October. And from then on every year that begins a new century must be exactly divisible by 400 to be a leap year. Thus 2000 and 2400 are leap years; 1900, 2100 and 2200 are not.

The change was immediately adopted by all the Catholic countries, but the Protestant and Greek Orthodox countries would have none of it. The furore eventually died down. Norway and Denmark changed their minds in 1700. Great Britain and its colonies (including America) followed suit in 1752. The bill passed by Parliament states in part "... and that the Natural Day next immediately following the said second Day of September shall be called, reckoned, and accounted to be the fourteenth Day of September omitting for that Time only the

Eleven intermediate Nominal Days of the Common Calendar. . .'. At the same time the beginning of the year (commencing with 1752) was changed from 25 March to 1 January. The 18 March 1751 Act of Parliament introducing these changes caused rioting in the streets and shouts of 'give us back our eleven days', the eleventh day having accumulated since Pope Gregory's proposal. Japan changed in 1873 and the Orthodox Church in Russia, Rumania and Greece in 1923.

Problems still exist, and Charles Kluepfel has investigated the accuracy of the gregorian calendar in *Sky and*

Telescope (64, 417; 1982). By AD 13000 the vernal equinox will again be occurring around 10 March. The 'new' gregorian calendar will then be just as much in error as the julian calendar was in 1582. Correcting the gregorian calendar appears, however, to be a bit premature. The length of the tropical year in terms of mean solar days depends on the Earth's spin rate. This is slowing down but the way in which this retardation is occurring is not accurately known. Also, the calculations depend on the determination of the tropical year at epoch 1900 made by the American astronomer Simon Newcomb. That needs to be recalculated. So all in all the gregorian calendar is at about the limits of its possible accuracy and the system will just need to be recalibrated every few thousand years. □

David W. Hughes is Lecturer in Physics and Astronomy at the University of Sheffield, Sheffield S3 7RH.



100 years ago

NEW OR RARE ANIMALS IN THE ZOOLOGICAL SOCIETY

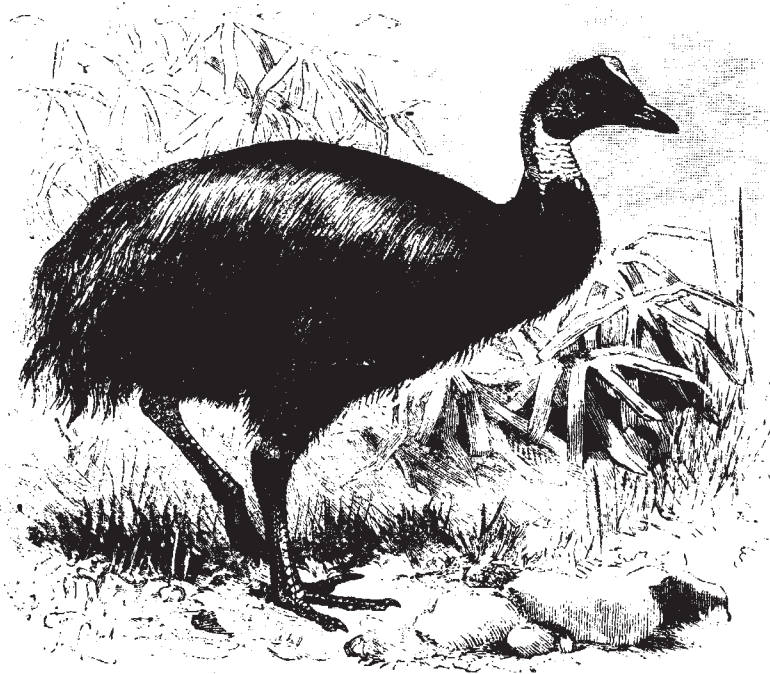
THE ONE-WATTLED CASSOWARY (*Casuarus uniappendiculatus*). The Cassowaries of the Moluccas and Papuan Islands, together with the Emeu of Australia, constitute a very well-marked division of the Struthionies, or ostrich-like order of birds, and occupy a large area of the Australian region. But while the Emeu (*Dromceus*) is spread over the whole of the Australian continent, the Cassowary is only met with in the northern parts of Queensland and in the peninsula of Cape York, and we

must cross Torres Straits into New Guinea and its adjoining islets before we arrive at the true metropolis of the Cassowaries. Here we shall find them scattered over the different islands to the number of nine, as indicated in Count Tommaso Salvadori's recent essay on the group, and but one, or at most two species being ever found exactly within the same area.

A characteristic of the Cassowaries is the large horny casque which covers the head, and is devoid of feathers. In one division of the genus this is much elevated and laterally compressed, in the other the casque is pyramidal in shape, and flattened cross-ways behind. The One-wattled Cassowary belongs to the second division, and is further distinguished by having (in common with its near ally *C. occipitalis*) but a single wattle in the middle of its throat.

This Cassowary was first made known to science in 1860 by Blyth, from an example brought alive to Calcutta, of which the exact origin was uncertain.

From *Nature* 27, 153, 14 December 1882.



Genetic code

Mitochondrial codes and evolution

from Thomas H. Jukes

It is a common assumption in evolution that all organisms in the terrestrial biosphere are descendants of a single ancestor, or a group of closely related ancestors. Indeed, a century ago, Pooh-Bah implied this when he said proudly that he could trace his ancestry back to a primeval protoplasmic atomic globule¹. Molecular evolutionists theorize that the ancestor used the present universal amino acid code². What preceded this code?

The code is embodied in transfer RNA (tRNA) molecules, each specific for an amino acid and containing a sequence of three nucleotides (an anticodon) that pairs in a reverse direction with three complementary nucleotides (a codon) in a messenger RNA (mRNA) strand. For example, anticodon GAU, for isoleucine, will pair with codon AUC, G pairing with C, A with U and U with A.

Some hints as to the nature of the earlier codes have come from mitochondria. Sequences of the complete DNA are known for human, bovine and mouse mitochondria³⁻⁵. Each contains genes for 22 tRNA molecules that suffice to code for

20 amino acids. The number of possible codons in mRNA is, by definition, 64, because this is the number of possible arrangements of A, C, G and U in groups of three. The 22 anticodons in mammalian mitochondria (MMt) pair with 60 codons, because there are four unpaired 'stop' codons for polypeptide chain termination in the MMt code. This compares with 55 anticodons in the universal code. The reduction in anticodons by MMt is achieved by using a single anticodon to pair with four different anticodons, for example, anticodon UAG pairs with CUU, CUC, CUA and CUG (leucine). The first anticodon base, U, pairs with U, C, A or G. Perhaps this is evidence of evolutionary retrogression towards a more primitive code, accompanying drastic evolutionary shortening of the mitochondrial genome.

A four-way pairing system would provide for only 15 amino acids in an archetypal code⁶, because there are only 16 possible anticodons starting with U, and one must be excluded to provide for chain termination. Four-way pairing is used eight times in the MMt code, and there are six

sets in which G pairs with U and C while U pairs with A and G (ref. 7 and the table), so that 20 amino acids are included. Stop codons UAA, UAG, AGA and AGG are provided by omitting their anticodons from the genome. There are two anticodons apiece for leucine and serine, which each have six codons.

The MMt code is simpler than the universal code, but not as simple as the proposed archetypal code⁶. This expanded to 20 amino acids by the standard evolutionary pathway of gene duplication followed by acquisition of a new function by one of the duplicated genes. A clue to this is in the MMt code. For example (see the table), UAA could have been the archetypal anticodon for codons UUN (N=U, C, A or G). It duplicated, following which one of the two duplicate tRNAs underwent a mutation to GAA in its anticodon. By a second mutation, this tRNA changed its aminoacylation site to couple with phenylalanine instead of leucine.

The suggestion that a change in aminoacylation can take place in evolution is supported by the finding that tRNA with anticodon UAG is aminoacylated by threonine in yeast Mt code and by leucine in all other known codes⁸, so that codons CUN code for threonine in yeast Mt and for leucine in other codes.

Further expansion was stopped by freezing of the code. This resulted from increasing complexity of living organisms

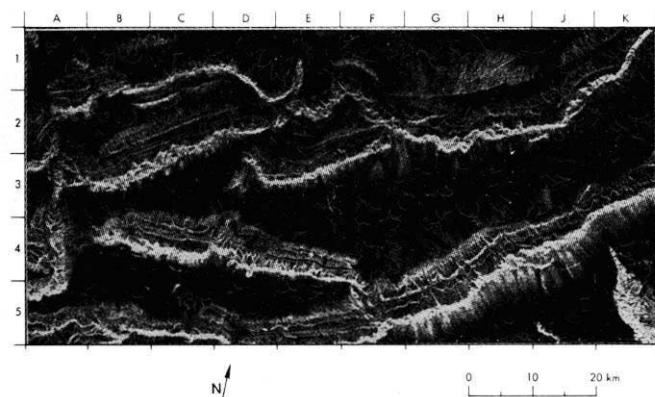


Fig. 1 Large-scale faulting can be seen in the radar image of the Kelpin Tagh uplift region of Northwestern China.

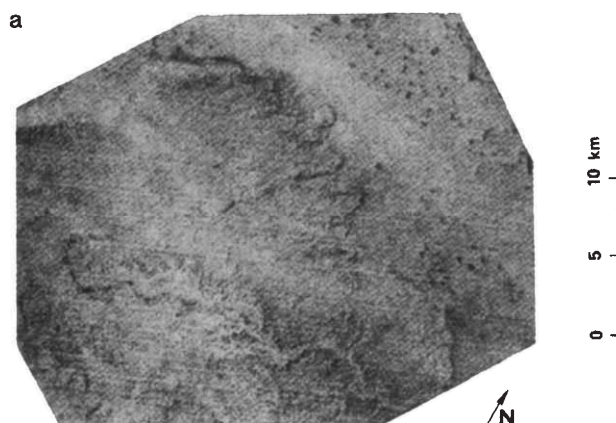


Fig. 2 (a) Landsat image and (b) radar image of the Telemzane crater region in Algeria.



AN imaging radar was one part of the first scientific payload to be taken into space by the shuttle *Columbia* in late 1981. High resolution images of the surface were obtained from analysis of time delays and Doppler shifts in the radar echo recorded in the spacecraft by optical holography (*Science* 218, 996; 1982). The radar image is a two dimensional record of the scattering properties of the surface — determined by shape, roughness and dielectric constant — and has a resolution of forty metres. The sensitivity of the radar to surface topography makes it particularly valuable in geological mapping. Faults, fractures and surface changes in roughness are easily picked out (Fig 1). Comparison with earlier landsat images (Fig 2a and Fig 2b) shows the advantage of the perpendicular radar illumination. In the landsat image, with solar illumination parallel to the fracture trends, the pattern of fractures is scarcely seen.

Anticodons in the mammalian mitochondrial code³⁻⁵. Anticodon GAA pairs with codons UUU and UUC, anticodon UAA with codons UUA and UUG, anticodon UAG with CUU, CUC, CUA and CUG, and so on, except that CAU pairs with AUA and AUG

GAA	Phe	UGA	Ser	GUA	Tyr	GCA	Cys
UAA	Leu	—	—	—	—	UCA	Trp
UAG	Leu	UGG	Pro	GUG	His	UCG	Arg
GAU	Ile	—	—	UUG	Gln	—	—
CAU	Met	UGU	Thr	GUU	Asn	GCU	Ser
UAC	Val	UGC	Ala	UUU	Lys	—	—
				GUC	Asp	UCC	Gly
				UUC	Glu		

to the point where changes in the code would be so disruptive of key proteins as to be lethal. Before this, organisms were small and primitive enough to endure such dislocations.

But the code froze only with respect to the number of amino acids in it. Duplication of tRNA genes continued and new anticodons appeared with higher codon-binding affinities, so that the fidelity of translation was improved. □

Thomas H. Jukes is at the Space Sciences Laboratory at the University of California, Berkeley, California 96720.

1. Gilbert, W.S. *The Mikado Act* 1 (1887).
2. Crick, F.H.C. *J. molec. Biol.* **38**, 367 (1968).
3. Anderson, S. *et al. Nature* **290**, 457 (1981).
4. Bibb, M.J. *et al. Cell* **26**, 167 (1981).
5. Anderson, S. *et al. J. molec. Biol.* **156**, 683 (1982).
6. Jukes, T.H. *Molecules and Evolution*, 69 (Columbia University Press, New York, 1966).
7. Crick, F.H.C. *J. molec. Biol.* **19**, 548 (1966).
8. Bonitz, S.B. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 3167 (1980).

New foods

Proving leaf protein's worth

from N.W. Pirie

THE second international conference on leaf protein* was held in Aurangabad in October 1982, some twelve years after the first had taken place in Coimbatore. In that interval, the potential of leaf protein has gained widespread recognition. The 92 papers submitted to the second conference came from 30 countries; scientists from 18 countries attended. As would be expected, most participants were from India: there is research on leaf protein, or a general interest in the subject, in 18 centres in that country.

Throughout the conference, small-scale production of leaf protein for local consumption was emphasized. An extraction unit was shown operating in a village near Aurangabad. Slides and a film of commercial leaf protein production for animal feed in France, New Zealand, Japan and US illustrated dramatically how little engineers know about the properties of leaf juice. Some equipment, originally designed for expressing oil from seeds, is used although it exerts unnecessarily intense pressure. To get protein-rich juice from pulped leaf it is inadvisable to use more than 2 or 3 kg cm⁻² (the pressure exerted between finger and thumb).

* Among Indian sources of finance for the conference were the University Grants Commission and the Indian National Science Academy. Assistance with travel grants came from the Commonwealth Foundation, the United Nations University and the Royal Society. R.N. Joshi and Narendra Singh made vigorous efforts to collect typescripts; publication of the proceedings of the conference is proposed.

Anything more than that is wasteful and, because leaf fibre becomes tightly packed and impenetrable, detrimental.

The consensus was that heat coagulation is the safest way to separate leaf protein from the juice. Because of the amount of energy needed to heat the juice, coagulation by acidification or anaerobic fermentation was proposed. Such methods introduce the risk that recoverable protein will be lost by proteolysis, that pheophorbide (a photosensitizing agent) will be formed and that harmful bacteria will proliferate. While some participants advocated relatively crude and hazardous methods of coagulation, others suggested refinements aimed at making partly decolorized protein. It is difficult to see what need such products would satisfy. They would be very much more expensive than leaf protein made simply by coagulation and washing, the processes used would be beyond the capacity of small-scale producers and useful lipids would be lost. By judicious use of counter-current heating the energy needed for heat coagulation can be diminished and, if a diesel rather than an electric motor is used for pulping and pressing, heat would be an abundant by-product.

A few participants objected, for political and pseudo-ethical reasons, to the use of leaf protein as a human food, but the majority could see no force in these

arguments. An immense range of leaf species is already eaten in different countries, though the amounts eaten are often less than is desirable. The processes of extraction, coagulation and separation will increase the potential range of species because most of the small molecules are removed, and leaf protein is exposed to no treatment that is not part of normal domestic cooking. Provided leaf species are chosen sensibly and processed skillfully, fears about possible toxicity are groundless.

Dried leaf protein has so far been used in all the feeding trials that have been published. The use of dried material is convenient, but drying is troublesome and diminishes the nutritive value of the protein. There was a useful paper on the preservation of moist leaf protein with three different cultures of lactobacillus, and a paper on the stability of β -carotene in moist leaf protein preserved with salt or acetic acid. The value of leaf protein as a food which supplies carotene (pro-vitamin A) in the finely divided lipid environment which facilitates absorption, was also stressed.

Besides leaf protein, fodder fractionation produces leaf fibre, from which part of the protein has been extracted, and the soluble components in a 'whey' which runs away from the coagulum of leaf protein. The former is economically the most important of the three products. Several papers dealt with its use as a ruminant feed, and with its conservation by drying or ensiling. There has so far been little work on the use of the 'whey' — it is, as a rule, returned to the land as fertilizer. However, the growth of yeasts and the production of penicillin on it were described. □

N.W. Pirie is in the Rothamsted Experimental Station, Harpenden, Herts AL5 2JQ.

Apology

We apologise to K.R. Kim and his collaborators H. Craig and Y. Horibe for publication of their graph of the methane anomalies above the Mariana Trough rift valley (*Call Alvin for hot science, News and Views* **215**, 214; 18 November 1982). The graph was circulated (and related results reported) at the 5th International Conference on Isotope Geology, Nikko but was not yet ready for general release. A full account of the work is currently in preparation.

Corrigendum

In the *News and Views* article 'Yet another opioid peptide' by Leslie L. Iversen (*Nature* **299**, 578; 1982), the statement that Corbett *et al.* (*Nature* **299**, 79; 1982) found dynorphin₁₋₈ to be "much more rapidly metabolized in brain than the full-length dynorphin" does not accurately describe what these authors reported. Their experiments with peptidase inhibitors suggested that this might be the case in the various peripheral tissue bioassay preparations used, but the idea that the same may hold true in brain was an inference which goes beyond presently available results. The correspondent apologizes for this inaccuracy.

Numerical dating of Precambrian–Cambrian boundary

G. S. Odin*, N. H. Gale†, B. Auvray‡, M. Bielski§, F. Doré||,
J.-R. Lancelot¶ & P. Pasteels*

* Département de Géologie dynamique, Université Pierre et Marie Curie, 75230 Paris Cédex 05, France

† Department of Geology and Mineralogy, Parks Road, Oxford OX1 3PR, UK

‡ Centre Armoricain d'étude structurale des socles, Campus de Beaulieu, 35042 Rennes Cédex, France

§ Institute of Earth Sciences, The Hebrew University, Jerusalem, Israel

|| Département de Géologie, Université, Esplanade de la Paix, 14032 Caen Cédex, France

¶ Laboratoire de Géochimie isotopique, U.S.T.L., 34060 Montpellier Cédex, France

* Eenheid Geochronologie, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium

Although the base of the Phanerozoic is an important date in the history of life and of the Earth's crust, few conclusive data have been available to estimate the numerical age of this boundary. Accurate and reliable modern data now exist, however, and those relevant modern radiometric dates are summarized in this review and are shown to be consistent with a considerably younger age for the Precambrian–Cambrian boundary than previously accepted.

THE Precambrian–Cambrian boundary is important because it is at the base of the Cambrian that the fossil record documents the first explosive event in the evolution of life. Although the Coelenterata and the Arthropoda are now thought to have made a hesitant début below the Tommotian it is only in lowest Cambrian¹ that marine life suddenly became widespread with the abundant appearance or radiation of diverse phyla of marine invertebrates—the Protozoa, Coelenterata, Archaeocyatha, Porifera, Bryozoa, Mollusca, Brachiopods, Arthropoda and Echinodermata. The numerical age of this boundary, together with an age of the Cambrian–Ordovician boundary, is necessary for reliable estimates of rates of evolutionary change within the Cambrian period. Previous estimates of the date of the base of the Cambrian, all based on rather uncertain data, range from 600 ± 20 Myr (ref. 2) through 595 Myr (ref. 3) to 570 Myr (refs 4, 5).

Disagreement in the estimate of the age of the Precambrian–Cambrian boundary may be linked with its uncertain definition. There is no agreement yet concerning either a lithostratigraphical (stratotype) nor a biostratigraphical (biozonal) definition of the boundary⁶.

However, recent reviews of the problem and possible solutions have been proposed^{7,8} in terms of outcrops which might be considered as the most satisfactory candidates for a lithostratigraphical definition. From these studies several points may be emphasized. The presence of skeletonized faunas is necessary to locate the Precambrian–Cambrian boundary—these faunas are virtually absent from the Precambrian but are present above it. Among the skeletonized faunas, most workers agree that the first occurrence of trilobites is not necessarily related to the basal Cambrian. According to the synthesis by Cowie⁸ small shelly fossils exist in Newfoundland, Central England, China and Siberia in a formation assigned to a pre-trilobite stage: the Tommotian. At present the first trilobites are generally thought to be characteristic of a second stage of the early Cambrian: the Atdabanian, and define valuable chronozones in it. However, the Tommotian Petrotsvet Formation on the Aldan in Siberia, contains both Tommotian small shelly fossils and Archaeocyathids and trilobites “generally similar to those from Morocco”⁸. In spite of the tentative use of archaeocyathid zones in Asian formations, no definite biozonation is available in the Tommotian, which renders correlations between these Asian outcrops and other basins very subjective for this transitional stage.

Thus, the Precambrian–Cambrian lithostratigraphical boundary still needs to be defined; however, the boundary is biostratigraphically located between a well known skeletonized

Atdabanian fauna with trilobites above and a more scarcely encountered Ediacarian type fauna of Vendian age, below. The term Ediacarian has been proposed¹ to denote a distinctive episode of uppermost Precambrian history but of Phanerozoic age. The transitional stage sometimes contains Tommotian faunas with specific worms and shelly fossils younger than those eventually observed in the underlying levels and also Archaeocyathids older than those observed above, while trilobites are generally absent.

The presently accepted estimate of the geochronological age of this boundary is based on very old data. In 1964 Cowie⁹ wrote that the “correlation (of the dated samples) with stratigraphical units was, even at the best, imprecise and often inconclusive”. From the 15 best dates that he selected, he proposed an age of 570 Myr for the base of the Cambrian. However, the dates used by Cowie must be recalculated using the now accepted International Congress of Sydney Conventional Constants (ICC) as has been done by Gale¹⁰. Most of the dates available above the boundary come from early Cambrian glauconies, the best of which are those of the Baltic Blue Clay, PTS 100: 522 ± 3 Myr ICC old (all analytical errors discussed here represent ± 2σ) and of the North American Murray Shale, PTS 183: 565 ± 30 Myr (ICC). Below the boundary and amongst various rocks, the late Precambrian Vire–Carolles granite, PTS 42: 563 ± 15 Myr (ICC) appeared especially useful to Cowie⁹. In 1971, Lambert³ included new Rb–Sr isochron dates: the two best of them coming from the late Precambrian Holyrood granites (PTS 352) which gives a recomputed age¹⁰ of 585 ± 15 Myr (ICC), and the Precambrian Northbridge granite–gneiss (PTS 353) which shows a recalculated age of 553 ± 10 Myr (ICC). In spite of these data, more recent reviews^{5,6,11} propose much older ages of 590 and 575 Myr (ICC) for the base of the Cambrian. The only new results proposed¹¹ are from Soviet glauconies for which, as mentioned already by Patchett *et al.*¹², no published analytical or geochemical data are available, and which, consequently, cannot be used. Moreover, Soviet authors have themselves questioned these dates¹.

Amongst the partially published data, the available results obtained in the Sinai already seem to be conclusive and the geological setting may be summarized as follows^{13–15}. The basement metamorphic rocks of the Arabian–Nubian massif, older than 600 Myr (Fig. 1), are intruded by a large granite batholith, dated at 580–600 Myr (refs 13, 15). These rocks are locally unconformably overlain by mainly alkaline volcanic rocks¹⁶ intruded in turn by small igneous bodies. An erosional phase follows leading to the formation of a peneplain on which the

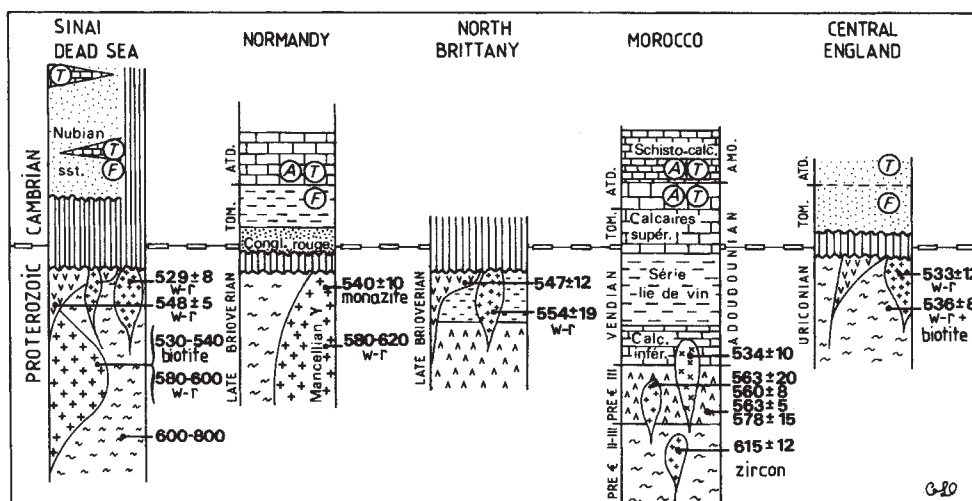


Fig. 1 Probable location of the Precambrian-Cambrian boundary in the best geochronologically documented areas. A, archaeocyathids; AMO, Amouslekian; ATD, Atdabanian; F, worms, skeletonized faunas and trace fossils; TOM, Tommotian; w-r, whole-rock Rb-Sr isochron ages; +, granite; ~, metamorphic rocks; v, rhyolite, ignimbrites.

fluviatile Nubian Sandstones were deposited. Close to the base of these unfossiliferous sediments, marine shales and dolomites are locally interbedded. They contain a trilobite fauna which has been correlated with that of Morocco¹⁷. In particular, species of the genus *Redlichops* and *Resserops* from the Dead Sea area have been correlated with trilobites of the Amouslekian and Timghitian substages of the Soussian stage; the first stage of the levels assumed to be early Cambrian in Morocco which is equivalent to the Atdabanian stage of Siberia. Consequently, the lower Nubian Sandstones, although not necessarily basal Cambrian, seem to represent a rather old Cambrian age.

Near Elat, in the North of the Massif, the Amram quartz porphyries are part of the volcanics unconformably overlying the granitic batholith; they unconformably underlie fossiliferous early Cambrian sediments. A nine-point Rb-Sr isochron age of 548 ± 5 Myr was calculated for these volcanics¹⁴.

In Southern Sinai, the peneplain has been destroyed by erosion but is estimated to have been 1–2 km above the surface on which one of the granites—the Mandar granite, is now exposed¹⁸. A five-point Rb-Sr isochron age of 529 ± 8 Myr was obtained for this granite. In addition, a regional thermal event affected the basement rocks as documented by mineral rejuvenation of both Rb-Sr and K-Ar apparent ages which yielded 530–540 Myr (refs 14, 15). No thermal effects have so far been found in the sediments above the peneplain. The post-peneplain early Cambrian sequence in the Arabian-Nubian Massif is therefore younger than 540 Myr.

The Vire-Carolles granite complex of Normandy (Massif Armoricain) had been dated previously. Several samples of this Mancelian granite of the Eastern Massif Armoricain have again been collected from different rock types¹⁹. At Villedieu especially, one of the samples comes from a few metres below the unconformity with the overlying unmetamorphosed Red Beds²⁰, a probably continental azoic formation (Fig. 1). The Red Beds underlie marine sediments which contain worms and trace fossils of arthropods (*Monomorphichnus*) and gastropods (*Taphrohelminthopsis*); these layers seem to represent the Tommotian stage^{21,22}. In the Carteret area, one may observe Atdabanian limestones with primitive trilobites (*Bigotina*) and archaeocyathids overlying levels with Tommotian or older faunas with heteractinellid sponges and hyolithids. The Red Beds locally underlie sediments very similar to these fossiliferous rocks. The Red Beds are therefore almost certainly of basal Cambrian age and even possibly latest Vendian (equivalent to the Nemakit Daldyn horizon). The most significant analytical results were obtained on four monazites of the granite which yield U-Pb ages falling between 531 and 551 Myr (refs 21, 23). Because of its frequent occurrence in Mancelian granites and euhedral shape in some cases, monazite may be regarded as a primary constituent of the pluton. The monazite age of 540 ± 10 Myr corresponds either to the age of

the granite emplacement and crystallization or to a later closure time which, however, would correspond to a rather high cooling temperature. Subsequently, uplift and erosion occurred with the deposition of the Red Beds following later. Thus the age of 540 ± 10 Myr corresponds to an upper limit for the age of sediments of Tommotian or older age.

In Brittany, in the northern Massif Armoricain, Rb-Sr results were obtained in the laboratory of Rennes on Upper Brioverian (Vendian) rocks. In this area, the Upper Brioverian sequence comprises a volcanic series followed by sediments characteristic of a slope. These rocks are intruded by granite and granodioritic plutons (the Mancelian intrusions) further to the east. Above this basement, an alkaline magmatism is developed, represented by an ignimbritic volcanism and by granites and microgranites. The lower Cambrian sequence outcropping in Normandy is unknown in Northern Brittany, but we have just shown that the Mancelian granites were followed by azoic Red Beds and higher by limestones with Atdabanian Archaeocyathids in the Carteret area²⁴. In spite of having no direct evidence, we consider that the intrusive rocks were emplaced before the deposition of the first sedimentary levels of Cambrian age observed elsewhere in the Massif.

The Mancelian intrusions of Normandy were dated on whole-rocks using the Rb-Sr method. The ages obtained lie between 617 ± 12 and 582 ± 40 Myr ICC²⁵. In northern Brittany, the alkaline magmatic rocks were sampled in the granite of Porz-Scarff (five samples), in the microgranite of Loguivy (six samples) and in the ignimbrites of Lézardrieux (five samples). The Rb-Sr analytical results may be arranged in two separate isochrons²⁶. The higher one, comprising all the data coming from granite and microgranite whole-rocks, reveals an age of 554 ± 19 Myr with an initial $^{87}\text{Sr}/^{86}\text{Sr}$ ratio of 0.7096 ± 0.0012 ; the lower one may be drawn using the data for the five ignimbritic whole-rocks and leads to an age of 547 ± 12 Myr with an initial Sr ratio of 0.7064 ± 0.0007 . Note the consistency between this age obtained on acid volcanics and the age coming from the granites; this confirms the lack of a later thermal event, and determines a maximum age for the Lower Cambrian sediments.

As emphasized above, the first conclusive modern results regarding the Precambrian-Cambrian boundary were obtained in Adoudounian (latest Precambrian) sediments (Fig. 1) intruded by a syenite in Morocco²⁷. These results were complemented by later work on older rocks for which we will give some details to show the coherence of the series of measurements²². In the Anti-Atlas area, the Precambrian II/Precambrian III boundary was dated at 615 Myr on zircons from the Bleida granodiorite²⁸. Also in the Anti-Atlas area, the Ouarzazate volcanic series (Precambrian III) is intruded by granites which yielded a U-Pb zircon age of 563 ± 20 Myr (ref. 29). Further to the south-west, in the Ifni area, Charlot obtained an age of 560 ± 8 Myr (ref. 30) on an equivalent intrusive pluton.

The volcanic series itself was first dated at 578 ± 15 Myr on zircons: rhyolite of the Bou Ourhioul in the Haut-Atlas²⁹. Another U-Pb zircon age of 563 ± 5 Myr has been obtained recently in Rennes on the ignimbrite of equivalent age at Tiouine³¹. Finally, the most significant result was obtained in the Adoudounian sedimentary rocks overlying these Ouarzazate volcanics. The Calcaire Inférieur is intruded by a syenite which revealed an U-Pb zircon age of 534 ± 10 Myr (ref. 27). Above the Calcaire Inférieur, lie the *Série lie de Vin*; both series are from the Adoudounian, a local stage considered as Vendian in age. Above them is the Calcaire Supérieur located in the Tommotian by Cowie⁸ because they contain the first trilobites (*Eofallotaspis*)³² and *Archaeocyathids* at their top. These stratigraphical data clearly show that the dated syenite gives a maximum age for the earliest Cambrian.

Finally, the recent results from the UK are also very conclusive. In the Wrekin area, samples were collected from the Ercall granophyre and the metamorphosed Rushton Schist (Fig. 1). The formations are disconformably overlain by shallow water sedimentary quartzites above which lie the Lower Comley Sandstones which contain brachiopods and hyolithids and may equate with the 'Sub-Holmia' zone of Tommotian stage. These overlying beds have yielded diagnostic early Cambrian trilobites in their upper part³³; these trilobites are roughly equivalent to those of the 'Callavia-Holmia' zone.

The Rushton Schist, regionally metamorphosed in amphibolite facies, and the Ercall granophyre must both have been cold when covered by the Lower Cambrian Sea. They are therefore unlikely then to have been affected by any isotopic resetting process, and the metamorphism of the Rushton Schist

is regarded as contemporaneous with the granophyre because no important younger intrusions are yet identified in the area. The eight-point whole-rock Rb-Sr isochron age obtained for the Ercall granophyre is 533 ± 12 Myr; the Rushton Schist yields a Rb-Sr biotite whole-rock cooling age of 536 ± 8 Myr (ref. 12). It is not accurately known where, inside the lower Cambrian, the unconformity is stratigraphically located; but it is not far from the base of the Cambrian and the dates obtained are maximum ages for the Cambrian marine transgression in the Wrekin area of the southern UK. Correlation on a world-wide basis will remain uncertain until there has been more work on the palaeontological zonation and nomenclature of Vendian, Tommotian and Lower Cambrian rocks.

According to the above evidence the age of the Precambrian-Cambrian boundary in the Massif Armoricain, Morocco, South Britain and the Arabian-Nubian Massif is as young as 530 Myr. We consider, however, that, because of the stratigraphical uncertainties in the available data, it is better to propose that the time interval 520–540 Myr almost certainly includes the boundary. This date is not incompatible with the data available to earlier reviewers if a proper estimate of stratigraphical, geochronological and analytical uncertainties is taken into account. It is indeed surprising that, in view of the evidence available before 1978, estimates for the boundary of the order of 570–590 Myr have continued to be made. Finally, although our proposed date for the Precambrian-Cambrian boundary may embarrass those who favour a date of about 519 Myr for the base of the Tremadoc³⁴, it is not inconsistent with the date of 495 ± 10 Myr for that boundary suggested by Odin and Gale^{35,36}.

1. Cloud, P. & Glaessner, M. F. *Science* **217**, 783–792 (1982).
2. Holmes, A. *Trans. Edin. geol. Soc.* **17**, 183–216 (1959).
3. Lambert, R. St. J. *geol. Soc. Lond. Spec. Publ.* **5**, 9–31 (1971).
4. Harland, W. B., Smith, A. G. & Wilcock, B. (eds) *Q. J. geol. Soc. Lond.* **120S** (1964).
5. Armstrong, R. L. *Am. Ass. petrol. Geol. Spec. Publ.* **6**, 73–91 (1978).
6. Cowie, J. W. & Glaessner, M. F. *Earth Sci. Rev.* **11**, 209–251 (1975).
7. Rozanov, A. Yu & Sokolov, B. S. *Geol. Mag.* **117**, 23–27 (1980).
8. Cowie, J. W. *Precamb. Res.* **15**, 199–206 (1981).
9. Cowie, J. W. *Q. J. geol. Soc. Lond.* **120**, 255–258 (1964).
10. Gale, N. H. in *Numerical Dating in Stratigraphy* (ed. Odin, G. S.) Ch. 25, 467–486 (Wiley, Chichester, 1982).
11. Cowie, J. W. & Cribb, S. J. *Am. Ass. petrol. Geol. Spec. Publ.* **6**, 355–362 (1978).
12. Patchett, P. J., Gale, N. H., Goodwin, R. & Humm, M. J. *J. geol. Soc. Lond.* **137**, 649–656 (1980).
13. Bielski, M. in *Numerical Dating in Stratigraphy* (ed. Odin, G. S.) Abstr. NDS 251, 943–948 (Wiley, Chichester, 1982).
14. Steinitz, G., Bielski, M. & Starinski, A. *ECOG VII, Jerusalem, Abstr.* (1981).
15. Bielski, M., Jäger, E. & Steinitz, G. *Contr. Miner. Petrol.* **70**, 159–165 (1969).
16. Bentor, Y. K. *Bull. Res. Coun. Israel*, **10G**, 19–63 (1961).
17. Parnes, A. *Israel. J. Earth Sci.* **20**, 179–205 (1971).
18. Kohn, B. P. & Eyal, M. *Earth planet. Sci. Lett.* **52**, 129–141 (1981).
19. Jonin, M. & Vidal, Ph. *Can. J. Earth Sci.* **12**, 920–927 (1975).
20. Chauris, L. *C. r. heb. Séanc. Acad. Sci., Paris* **242**, 3092–3094 (1959).
21. Pasteris, P. & Doré, F. in *Numerical Dating in Stratigraphy* (ed. Odin, G. S.) Abstr. NDS 121, 784–790 (Wiley, Chichester, 1982).
22. Cahen, L. & Doré, F. *Q. J. geol. Soc. Lond.* **120S**, 300–302 (1964).
23. Pasteris, P. *Eclogae geol. Helv.* **63**, 231–237 (1970).
24. Doré, F. *Bull. Soc. geol. Fr.* **14**, 79–93 (1972).
25. Vidal Ph. *Mem. Soc. geol. miner. Bretagne* **21**, 162 p., (1980).
26. Auvray, B. in *Numerical Dating in Stratigraphy* (ed. Odin, G. S.) Abstr. NDS 249, 938–941 (Wiley, Chichester, 1982).
27. Ducrot, J. & Lancelot, J.-R. *Can. J. Earth Sci.* **14**, 2771–2777, (1977).
28. Ducrot, J. *Bull. Soc. geol. Fr.* **21**, 495–499 (1979).
29. Juery, A. thesis Univ. Paris (1976).
30. Charlot, R. thesis Univ. Rennes (1978).
31. Mifdal, A., Peucat, J. J. & Charlot, R. in *Numerical Dating in Stratigraphy* (ed. Odin, G. S.) Abstr. NDS 250, 941–943 (Wiley, Chichester, 1982).
32. Szduz, K. *Geol. Mag.* **115**, 83–94 (1978).
33. Cowie, J. W., Rushton, A. W. & Stubblefield, C. J. *Spec. Rep. geol. Soc. Lond.* **2**, 100 (1972).
34. McKerrow, W. S., Lambert, R. St. J. & Chamberlain, V. E. *Earth planet. Sci. Lett.* **51**, 1–8 (1980).
35. Odin, G. S. & Gale, N. H. *C. r. heb. Séanc. Acad. Sci., Paris* **294**, 453–456 (1982).
36. Odin, G. S. (ed.) *Numerical Dating in Stratigraphy* (Wiley, Chichester, 1982).

ARTICLES

The Smoking Hills: natural acidification of an aquatic ecosystem

Magda Havas & Thomas C. Hutchinson

Department of Botany and Institute for Environmental Studies, University of Toronto, Toronto, Canada M5S 1A1

Spontaneous burning of bituminous shales at the Smoking Hills in the Canadian Arctic has produced intense acidic fumigations and strongly influenced the local tundra. The burns are of great antiquity. In an area of typically alkaline ponds with pH above 8.0, ponds within the fumigation zone have been acidified below pH 2.0. Elevated concentrations of metals (aluminium, iron, zinc, nickel, manganese and cadmium) occur in these acidic ponds. Soils and sediments have also been chemically altered. The biota in the acidic ponds are characteristic of acidic environments worldwide, in contrast to the typically Arctic biota in adjacent alkaline ponds.

ECOSYSTEMS which have been subjected to long-term airborne acidification are rare. Studies of such systems are of interest, not only in assessing the biological and chemical

responses to long-term acid inputs, but also because they provide useful and possible predictive information as to the long-term consequences of regional acid precipitation.

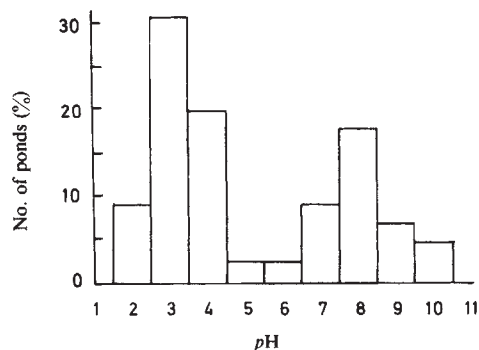


Fig. 1 Frequency distribution of pond water pH for 46 tundra ponds at the Smoking Hills.

Such an ecosystem occurs in the western Canadian Arctic at the Smoking Hills area of Cape Bathurst, North West Territories ($70^{\circ}14'N$, $127^{\circ}10'W$). Here, along 30 km of sea cliffs, exposed bituminous shales are spontaneously burning. Dense clouds of white smoke are emitted from several fumaroles. The smoke, which consists of sulphur dioxide, sulphuric acid aerosol, and steam, is carried inland across the tundra and occasionally out over the sea-ice and open waters of Franklin Bay. New fumigations arise from time to time, especially in the spring, as cliffs are eroded and slide to the sea, exposing fresh shale deposits.

The deposits are of Upper Cretaceous age, and are comprised of black shales interbedded with yellow bands of jarosite ($KFe_3(SO_4)_2(OH)_6$). Spontaneous combustion occurs as a result of finely divided sulphur coming into contact with bitumen and producing an exothermic reaction, much as it does in colliery spoil heaps. The burned material lies in huge brick-red and white mounds on the beaches. The jarosite and the shales emit sulphur dioxide when heated.

The burns at the Smoking Hills were first reported by nineteenth century explorers seeking the North-West Passage. Sir John Franklin's expedition examined the phenomenon and the physician-naturalist Richardson¹ described the sulphurous nature of the emissions. He suggested the sulphur-bitumen exothermic reaction, and noted the 'sour' taste of pools and ponds in the vicinity. The enormous quantity of burnt material in the cliffs and on the shore, in relation to the slow rate of burning observed over the period 1975–81, together with widespread areas of dead vegetation and burnt sea cliffs in regions where burns no longer occur, strongly suggest an antiquity of several hundred to several thousand years.

Although spontaneous burning of bituminous shales has been reported from other areas^{2–4}, their impact on the local environment has not been studied. We now examine the effects of the acid fumigations on the chemistry of local tundra ponds and the occurrence of biota in these chemically altered waters.

Studies were done during the short summer growing season from 1975 to 1981. Chemical analyses were carried out in Toronto using standard techniques including a turbidimetric method for sulphate, IR gas analysis for total and inorganic carbon, atomic absorption spectrophotometry and neutron activation analysis for major cations⁵.

Site description

The Smoking Hills is in a region of continuous permafrost and has a *Dryas integrifolia*-dominated flora. The summer active layer is 20–40 cm deep. Soils are derived from calcareous marls and are extremely well buffered. Such a region might be assumed to be virtually immune from the damaging effects of acid precipitation. But this is not the case. The striking feature of the Smoking Hills study is that effects have been both severe and extensive.

Pond water chemistry

The pH of 46 pond waters, sampled within a 20 km² area centred on the major fumigation, showed a wide range of values

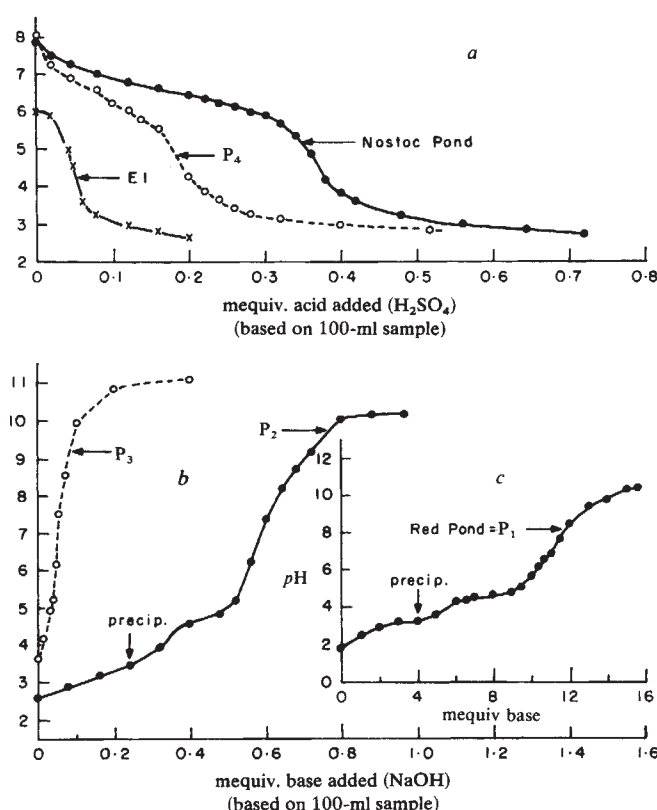


Fig. 2 Acid-base titration curves for Smoking Hills ponds. *a*, Three alkaline ponds; *b* and *c*, three acidic ponds.

(1.8–9.7). The acidic ponds were closest to the burning cliffs. As a result of direct acidic fumigations and drainage from acidified watersheds, ponds which presumably were once alkaline and well buffered (0.6 – 3.7 mequiv l^{-1} or 30 – 185 mg l^{-1} alkalinity) have been titrated to extremes of acidity (6.4 – 120 mequiv l^{-1} or 320 – $6,000$ mg l^{-1} acidity) and to pHs as low as 1.8. The pH range of ponds at the Smoking Hills is clearly much broader than that encountered in the granitic Precambrian Shield areas of eastern North America and Scandinavia which are presently subjected to acid precipitation^{6–8}. The severe effects must be ascribed to the long period of exposure and to the intensity of acidification. We have recorded sulphur dioxide levels >1 p.p.m. over several days near the cliff top and pH of summer rainstorms falling through the plume as low as 2.0. Continuous summer and winter deposition of sulphur dioxide, sulphate and acidic aerosol add to the impact.

The low pH of the most acidic ponds is more reminiscent of acidic hotsprings⁹, acid mine drainage ponds and streams^{10–13}, or of the strongly acidic volcanic lakes described from Japan^{14–16}, than it is of industrially-induced acid rain effects.

A further feature of pond-water chemistry is the marked bimodal frequency distribution of pond pH (Fig. 1). That only 2 of the 46 ponds were in the pH range 4.5–6.5 is explained by the buffering systems present in these ponds. The alkaline ponds are buffered by the bicarbonate-carbon dioxide system and the acidic ponds by an aluminium-iron system (Fig. 2). An intermediate buffering system is absent. This bimodal pH frequency distribution has also been noted in areas exposed to regional acid precipitation in Scandinavia, Ontario, Canada and in the Adirondacks of the USA^{17–19}.

Alkaline ponds differ markedly in their acid-neutralizing capacity and this correlates with their distance from the fumigations. Titration of alkaline pond water with sulphuric acid shows that once the bicarbonate buffering is overcome, the pH falls rapidly with continued acid additions (Fig. 2*a*). Interestingly, when alkaline pond water was exposed to Smoking Hills fumigations in plastic containers with and without their sediments, the sediments were found to be especially strongly buffered and

could reverse water column acidification quite rapidly when fumigations were excluded again.

Titration of acidic pond water (P_1 and P_2) with sodium hydroxide showed that as the pH was raised to 3.5, an orange-red precipitate consisting primarily of ferric hydroxide was observed (Fig. 2b, c). The pH then changed rapidly until another well buffered plateau between pH 4.0 and 4.5 was reached; this related to aluminium hydroxide buffering. The concentrations of metals in P_3 were considerably lower than in either P_1 or P_2 due to the higher pH and to the complete covering of the bottom sediments by vegetation. As a result, P_3 had a lower base-neutralizing capacity than P_1 and P_2 (Fig. 2b and Table 1).

The frequency distribution of pH values (Fig. 1) can be taken to represent the time progression in pH change experienced by an individual pond as it is fumigated. Organisms initially present in the ponds are faced with rapid changes in water chemistry once the pH falls below 6.5. The time for physiological and evolutionary adaptation during this accelerated phase is short. Most species have not been able to make this transition.

The concentrations of several potentially toxic elements increased with acidity (Table 1). Iron, aluminium and manganese concentrations were especially high. Levels of zinc, nickel and cadmium were also elevated in the most acidic waters. In the two ponds (P_1 and RP) closest to the fumigation, uranium, vanadium, cobalt and arsenic concentrations were in the mg l^{-1} range. Biota, in order to survive, need to be not only acid-tolerant, but also tolerant of a mixed-soup of potentially toxic elements, each occurring at high concentrations.

Aluminium (at 0.3 mg l^{-1}) is believed to be the major factor responsible for the loss of fish in Adirondack lakes²⁰. In the Smoking Hills acid ponds, levels of aluminium $>800 \text{ mg l}^{-1}$ occur. Iron, which is predominantly in the toxic ferric form in these highly acidic waters, occurs at $2,600 \text{ mg l}^{-1}$ (Table 1).

Calcium levels are high in the alkaline waters, as one might expect in regions of calcareous marls. However, the acidic conditions created by fumigations also have mobilized bases from the soils and sediments (Table 1 and Fig. 4). Calcium, magnesium and barium concentrations in pond P_1 were 370, 500 and 230 mg l^{-1} respectively. Elevated calcium levels in lakes receiving acid precipitation have been shown by Henriksen^{21,22}, although at considerably lower concentrations than observed at the Smoking Hills.

Sulphate concentrations of several thousand mg l^{-1} occur in the most acidic water at the Smoking Hills. Some of the alkaline ponds also contain high sulphate concentrations. All ponds within the 20 km^2 area are exposed to some fumigation and this partly accounts for the high sulphate levels. Watershed sources, including mineralization of organic sulphur compounds and inputs of marine aerosols, also contribute.

A plot of the atmospheric hydrogen ion inputs against sulphate inputs in the study area, using data from dustfall-rainfall buckets, produces a highly significant correlation ($r = 0.996$)²³. The theoretical dissociation line for sulphuric acid almost overlaps when plotted on the same scale. Clearly, at the Smoking Hills, all of the atmospheric acidic input to the ponds can be ascribed to sulphur sources. Unlike most regional acidification

Table 1 Chemical analyses of filtered ($0.45 \mu\text{m}$) pond water from the Smoking Hills

a, Results for a total of 45 ponds*

pH range	1.5–2.5	2.5–3.5	3.5–4.5	4.5–5.5	5.5–6.5	6.5–7.5	7.5–8.5	8.5–9.5	9.5–10.5
Sample size	4	14	9	1	1	4	8	3	2
Al	270	5.5	1.1	<0.6	<0.2	<0.8	<0.7	<0.2	<0.2
Fe	500	18	1.2	0.5	0.2	<0.04	0.1	<0.04	0.2
Mn	61	155	3.6	23	1.8	0.7	<0.5	<0.2	1.2
Zn	14	0.45	0.12	0.03	0.05	0.08	0.04	<0.03	5.3
Ni	6.3	0.21	0.04	0.04	0.06	0.02	0.004	0.007	0.01
Cd	0.520	0.022	0.011	0.004	0.012	0.001	0.003	<0.001	<0.001
Ca	301	157	44	182	249	90	49	31	16
SO_4	8,200	890	156	813	713	360	106	34	29
Cl	59	88	27	83	13	46	55	21	17

b, A more detailed analysis of seven of the ponds in a

Parameter	Unit	P_1	RP	P_2	P_3	P_4	P_5	P_6
pH	$-\log(\text{H}^+)$	1.8	2.0	2.8	3.6	8.2	8.1	8.1
Acidity	$\text{mg CaCO}_3 \text{ l}^{-1}$	5,800	6,000	320	30	0	0	0
Alkalinity	$\text{mg CaCO}_3 \text{ l}^{-1}$	0	0	0	0	95	105	—
Conductivity	μS	12,500	13,800	100	370	2,400	750	1,200
Inorganic C	mg l^{-1}	—	—	2.3	5.0	24	27	—
Organic C	mg l^{-1}	—	—	4.6	4.8	37	21	—
$\text{PO}_4\text{-P}^\dagger$	mg l^{-1}	—	—	0.035	0.028	0.022	0.01	—
$\text{NO}_3\text{-N}^\ddagger$	mg l^{-1}	—	—	12	1.3	0.5	—	—
SO_4	mg l^{-1}	16,000	17,000	380	160	890	110	280
Cl	mg l^{-1}	6.0	12	18	14	320	140	145
Na	mg l^{-1}	550	—	21	13	190	56	60
Mg	mg l^{-1}	500	650	22	13	140	42	50
Ca	mg l^{-1}	370	360	45	49	250	61	87
Ba	mg l^{-1}	230	40	<1.8	<1.6	<2.3	<1.8	<1.8
Al	mg l^{-1}	590	810	18	3.8	<0.6	<0.3	<0.3
Mn	mg l^{-1}	64	69	2.8	3.5	<0.3	<0.2	<0.2
Fe	mg l^{-1}	2,600	—	32	1.1	0.12	0.04	—
Ni	mg l^{-1}	22	80	0.19	0.08	0.03	0.03	—
V	mg l^{-1}	4.2	5.1	<0.01	<0.01	<0.02	<0.01	<0.02
U	mg l^{-1}	1.3	1.2	0.04	<0.01	<0.01	<0.01	<0.01
Co	mg l^{-1}	2.2	1.9	0.06	<0.03	<0.04	<0.03	<0.03
As	mg l^{-1}	2.3	—	<0.01	<0.01	<0.01	<0.01	—
Distance§	m	40	60	190	670	2,640	4,430	4,280

* All data in a are expressed in mg l^{-1} .

† Soluble reactive phosphate.

‡ Value may be high due to chemical interference.

§ From major fumigation.

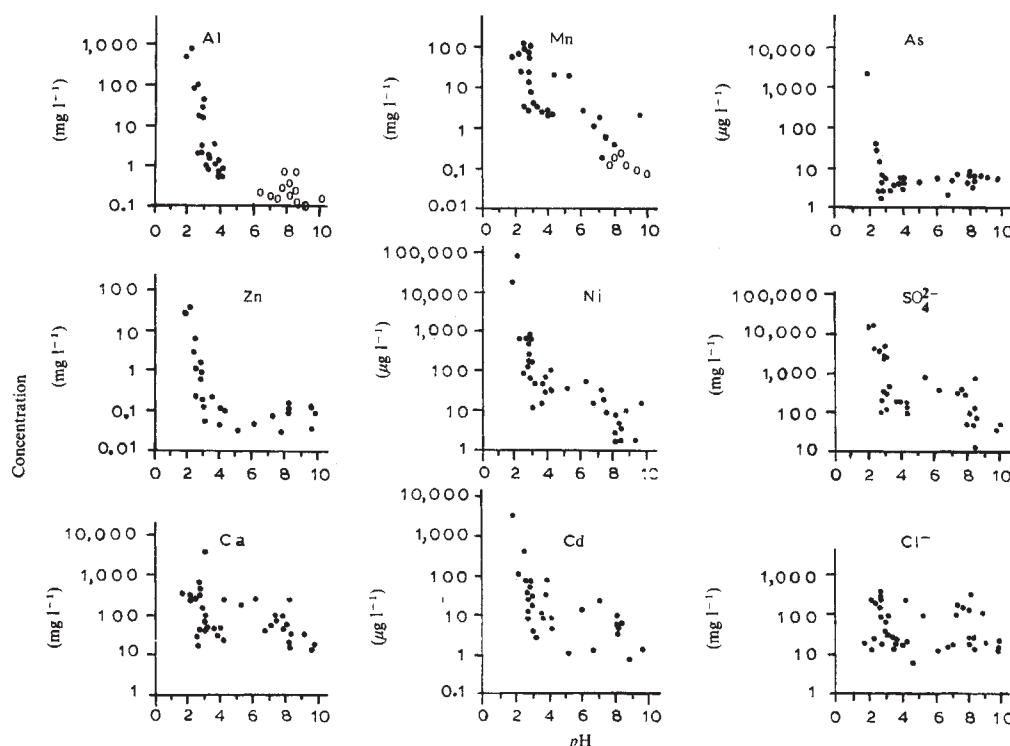


Fig. 3 The relationship between pH and the concentration of selected elements in Smoking Hills pond water over the pH range 1.8–9.7. ○, Below limit of detection.

problems, we do not have an additional atmospheric nitrogen oxide source to consider.

Solubility relations between the pH of pond waters and their elemental concentrations are illustrated in Fig. 3. The exponential increase in aluminium below pH 4.0 is apparent. Iron (not shown) and zinc have similar patterns. Small changes of pH in the region of pH 4.0 clearly will have substantial consequences to metal bioavailability and toxicity. Concentrations of manganese, nickel and cadmium also increase substantially at low pH but have greater scatter at any one pH. Vanadium (not shown) and arsenic, which presumably exist as anions (vanadate and arsenate) rather than cations, are elevated only in the most acidic ponds below pH 2.8. Calcium, sodium (not shown) and chloride do not show a pH-dependent pattern. However, a plot of the calcium-to-chloride ratio, which standardizes for marine-derived aerosols, does show an increase at lower pH. Combined with soil²³ and sediment chemistry (Fig. 4), these data strongly suggest calcium leaching.

Soil and sediment chemistry

There is a considerable difference in the chemistry of soils and sediments within the fumigated area compared with the non-fumigated or reference area. Conversely, soils and sediments collected from the same area are remarkably similar both in terms of the actual concentrations and patterns observed through the profile. The effects of the fumigation on soils and sediments have been severe. In the non-fumigated area, soil pH is >7.5 compared with 3.0 in the fumigated area. Surface sediments in acidified ponds are even more acidic than the soils. Several elements (calcium and manganese) have been leached from the soils and sediments, presumably by their solubilization in the acidic conditions (Fig. 4). Calcium, for example, is reduced from 4–5% in alkaline-pond sediments to <0.25% in acid-pond sediments. Chloride shows a similar pattern, although not necessarily a similar mechanism. In contrast, other elements (aluminium, barium, uranium, vanadium and sodium) are increased in sediments and soils within the fumigated area compared with the reference area. Deposition from the smoke plumes explains some, if not all, of this accumulation. Changes in oxidation–reduction conditions in cold wet soils are also

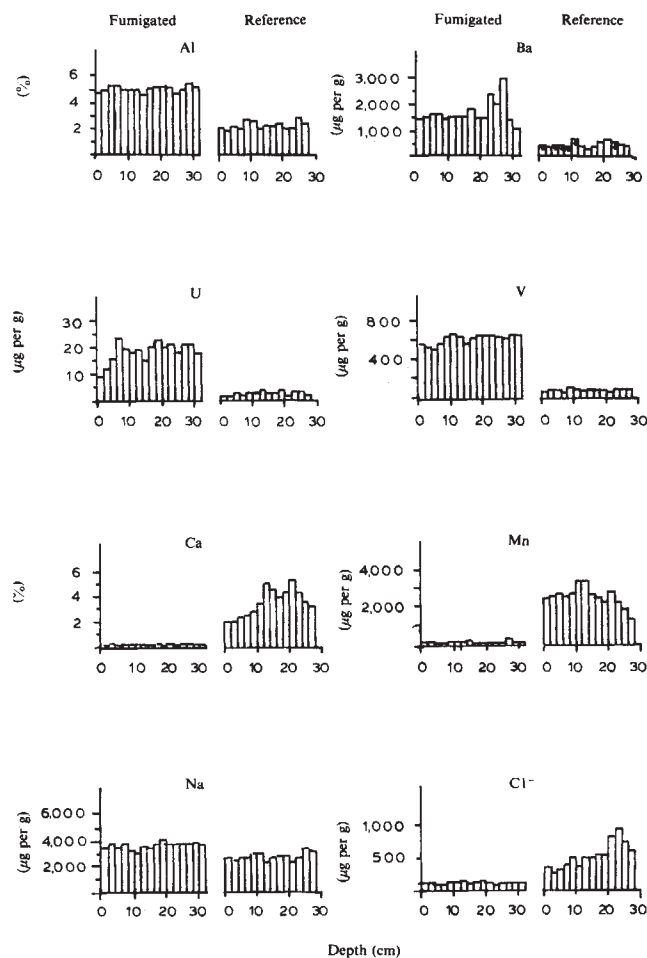


Fig. 4 Elemental composition of pond sediment cores from a fumigated pond (pH 2.8) and from a reference or non-fumigated pond (pH 8.2). Arrows indicate limit of detection.

likely to influence the relative solubility and therefore the mobility of certain elements.

Leaching of the bases calcium and magnesium, has been one major concern expressed about the potential impact of acid precipitation on forest soils²⁴. From the Smoking Hills data, it seems that even soils and sediments of very high base status can be depleted given a sufficient time and intensity of exposure to acid inputs.

Flora

One of the primary interests in examining the Smoking Hills ponds was to determine what biota, if any, were able to survive in the severe conditions of very low pH and high metal concentrations. No matter how severe the conditions, no pond was devoid of life. Algae were the most persistent. Full details are given elsewhere²⁵. In four acidic ponds (pH 1.8–3.6), sampled between 1975 and 1979, 14 species from four classes (Chlorophyceae, Euglenophyceae, Bacillariophyceae and Cryptophyceae) were identified. *Euglena mutabilis* was the dominant alga in the most acidic and metal-contaminated waters (P₁ and RP). *Nitzschia communis* and *Eunotia glacialis* were also present. Others common in the acidic waters were *Chlamydomonas acidophila*, *Eunotia arcus* and *Nitzschia palea*. These species are not so much typical of tundra ponds as of very acidic habitats worldwide^{26,27}. In laboratory studies³⁴, field isolates of *Euglena* were always associated with a yeast, *Cryptococcus* sp., in a mutualistic association which enhanced acid and metal tolerance of both species.

In alkaline ponds, 90 algal species belonging to nine classes were identified. The algal flora was typical of tundra ponds dominated by small flagellates, especially the cryptomonad *Chroomonas minuta*. Diatoms such as *Amphora*, *Navicula* and *Surirella* species accounted for 70–90% of the biomass of the periphyton.

In comparing alkaline to increasingly acidic ponds, the importance of planktonic as opposed to benthic biota declined. The bottom of the acidic ponds (down to pH 2.8) was covered with a dense mat of bryophytes, largely composed of the mosses *Drepanocladus exannulatus* and *Leptodictyum riparium*. This latter species has been reported in Japanese volcanic lakes of pH 1.6 (ref. 28), and in lakes around the Sudbury smelters at pH 4.0 (ref. 29). In Smoking Hills ponds below pH 2.8, these benthic mosses were absent.

Fauna

Zooplankton diversity is low in Arctic ponds compared with temperate lakes. In the alkaline ponds at the Smoking Hills, only 11 species were found. Nevertheless, density and biomass were high. The commonest species were *Daphnia middendorffiana*, *Diaptomus arcticus*, the fairy shrimp *Branchinecta paludosa* and the tadpole shrimp *Lepidurus arcticus*. The fairy shrimps swarm at the surface in summer and are fed on by Arctic terns and Arctic loons. Since the ponds freeze solid in winter, fish are absent.

Planktonic crustaceans were absent from all acidic ponds although a very few cyclopoid copepods occurred in some ponds at pH 3.5. The sensitivity of these Smoking Hills crustaceans to acid waters has been demonstrated experimentally^{30,31}. In acidic waters, crustaceans were replaced by rotifers. *Brachionus urceolaris* was particularly abundant (up to 41,000 individuals m⁻³). Again, this is a species known to occur in acid waters from Japan¹⁵. Rotifers are also reported to predominate in Norwegian lakes at low pH (ref. 6).

The larvae of only a few insect species were found in the ponds at the Smoking Hills. A caddis fly larva *Limnephilus pallens* (Trichoptera) and a chironomid larva *Orthocladus consobrinus* (Diptera) were both common in the alkaline ponds. Neither occurred in the acidic waters. Instead, a red chironomid larva, *Chironomus riparius*, was abundant in the acid ponds living amongst the benthic bryophytes. The occurrence of red chironomids has been reported from other acidic, metal-contaminated environments, such as acid-mine drainage streams³². It has been suggested that the presence of haemoglobin increases the internal buffering of body fluids and thereby increases tolerance to low pH³³. Crane fly larvae and a few oligochaetes were also found. In ponds below pH 2.0, neither insect larvae nor rotifers were observed.

Conclusions

At the Smoking Hills spontaneous burning of bituminous shales over hundreds of years has produced strongly acidic sulphur fumigations, and has affected the vegetation, soils and ponds. Despite the originally well buffered nature of the soils and pond water, the duration and severity of the acidifications have caused profound changes in their chemistry. The ponds have been acidified to pHs as low as 1.8. The frequency distribution of pond water pH in the area is strongly bimodal, reflecting the two buffering systems present (bicarbonate-carbon dioxide and aluminium-iron). The number of species is substantially reduced in the acidic ponds, some of which contain extremely high concentrations of aluminium, iron, manganese and zinc, as well as potentially toxic levels of nickel, cadmium, uranium, vanadium and arsenic. The most acid- and metal-tolerant plants seem to be mosses and benthic-dwelling algae. The principal animals in the very acidic waters are rotifers and a red chironomid larva. The biota of the alkaline ponds are typically Arctic in character while those of the acidic ponds are characteristic of the biota described from strongly acid waters worldwide. They are not a subset of the much more diverse biota of the alkaline ponds but have invaded *de novo* presumably pre-adapted to the acidic conditions.

This work was supported by grants to T.C.H. from the National Science and Engineering Research Council of Canada (NSERC), the Fisheries and Marine Department of Environment Canada and by the NATO Eco-Sciences Panel. M.H. was supported by a NSERC scholarship. Logistic support in the field was provided by the Polar Continental Shelf Project of Energy, Mines and Resources.

Received 9 September; accepted 5 November 1982.

- Franklin, Sir J. *Narrative of a Second Expedition to the Shores of the Polar Sea: 1825–1827* (Murray, London, 1828).
- Selwyn, A. R. C. *Report on Explorations in British Columbia* (Geological Survey of Canada, 1877).
- Henderson, G. *Geol. Surv. Greenland Rep. 22*, Copenhagen (1969).
- Rosenkrantz, A. *Grønland 12*, 377–384 (1977).
- Havas, M. thesis, Univ. Toronto (1980).
- Almer, B., Dickson, W., Ekstrom, C., Hornstrom, E. & Miller, U. *Ambio 3*, 330–336 (1974).
- Gjessing, E. T., Hendriksen, A., Johannessen, M. & Wright, R. F. *SNSF Proj. Res. Rep. FR6/76* (1976).
- Dillon, P. J. et al. *J. Fish. Res. Bd Can.* **35**, 809–815 (1978).
- Brock, T. D. *Science 179*, 480–483 (1973).
- Lackey, J. B. *Public Hlth Rep. 54*, 740–746 (1939).
- Joseph, J. M. *Ohio J. Sci.* **53**, 123–127 (1953).
- Warner, R. W. *Ohio J. Sci.* **71**, 202–215 (1971).
- King, D. L., Simmler, J. J., Decker, C. S. & Ogg, C. W. *J. Water Pollut. Cont. Fed.* **46**, 2301–2315 (1974).
- Yoshimura, S. *Archiv. Hydr.* **26**, 197–202 (1933).
- Negoro, K. *Jap. J. Limnol.* **9**, 176–179 (1959).
- Ueno, M. *Verh. int. Ver. Limnol.* **13**, 217–226 (1958).
- Wright, R. F. & Gjessing, E. T. *Ambio 5*, 219–223 (1976).

- Conroy, N., Hawley, K. & Keller, W. *Rep. Water Resources Assessment, North-eastern Region* (Ministry of Environment, Ontario, 1978).
- Schofield, C. L. *Ambio 5*, 228–230 (1976).
- Baker, J. P. & Schofield, C. L. in *Proc. Int. Conf. Ecological Impact of Acid Precipitation*, Norway (eds Drablos, D. & Tøllan, A.) 292–293 (SNSF Project, 1980).
- Hendriksen, A. *Nature 278*, 542–545 (1979).
- Hendriksen, A. in *Proc. Int. Conf. Ecological Impact of Acid Precipitation*, Norway (eds Drablos, D. & Tøllan, A.) 68–74 (SNSF Project, 1980).
- Gizyn, W. I. thesis, Univ. Toronto (1980).
- Norton, S. A. in *Proc. 1st Int. Symp. Acid Precipitation Forest Ecosystem*, 711–718 (USDA Forest Service, General Tech. Rep. NE-23, 1976).
- Sheath, R. G., Havas, M., Hellebust, J. A. & Hutchinson, T. C. *Can. J. Bot.* **60**, 58–72 (1975).
- Cassin, P. E. *J. Phycol.* **10**, 439–447 (1974).
- Hargreaves, J. W., Lloyd, E. J. H. & Whitton, B. A. *Freshwat. Biol.* **5**, 563–571 (1975).
- Tamura, T. *Jap. J. Limnol.* **6**, 63–73 (1936).
- Gorham, E. & Gordon, A. G. *Can. J. Bot.* **41**, 371–378 (1963).
- Havas, M. & Hutchinson, T. C. *Can. J. Fish. Aquat. Sci.* **39**, 890–903 (1982).
- Havas, M. & Hutchinson, T. C. *Can. J. Zool.* (in the press).
- Lackey, J. B. *Publ. Hlth Rep.* **53**, 1499–1507 (1938).
- Jernelov, A., Nagell, B. & Svenson, A. *Holarctic Ecol.* **4**, 116–119 (1981).
- Hutchinson, T. C., Nakatsu, C. & Tam, D. *Int. Conf. Heavy Metals in the Environment*, Amsterdam, 300–304 (CEP Consultants, Edinburgh, 1981).

New evidence from Fraser Cave for glacial age man in south-west Tasmania

Kevin Kiernan*, Rhys Jones† & Don Ranson‡

* Department of Geography, University of Tasmania, Hobart, Tasmania

† Department of Prehistory, Research School of Pacific Studies, Australian National University, Canberra, Australia

‡ Archaeological Section, Tasmanian National Parks and Wildlife Service, Hobart, Tasmania

The discovery of a rich archaeological occupation site in the Franklin River valley of south-west Tasmania dated at 15–20 kyr BP is described. The stone tools found support the view that the Tasmanian industries were derived from contemporary mainland Australian ones when the two were still connected. Faunal remains from human hunting and cooking activities indicate a hunting strategy targetted at a few favoured species, especially the large wallaby. No extinct megafaunal species were present. During the height of the last ice age, these glacier edge hunters of southern Tasmania were then the most southerly humans on Earth.

ACCORDING to ethnographic sources of the early nineteenth century, the densely forested inland region of south-west Tasmania was not then occupied by Aborigines¹. G. A. Robinson, who between 1829 and 1834 combed the island to meet most of the surviving Aborigines and persuade them to enter Government settlements, said of the region from a vantage point on the Arthur Range to the south on 13 March 1830 that "there was not the least sign or appearance of natives or of any white man ever being in this part of the country. The natives that accompanied me assured me there was [sic] no natives ever went inland".² It was because of this very inaccessibility that a notorious maximum punishment convict station was established on a small island in Macquarie Harbour in 1825. However, comments from Goodwin, one of the few convicts to escape across the mountains from here, and other explorers suggest that there was at least seasonal Aboriginal use of tongues of grassy country on the Upper Gordon River to the east and around Frenchman's Cap to the north³. A radiocarbon date of 300 ± 150 yr BP (ANU 2787) from an open river bank site on the confluence of the Gordon and Denison rivers (Fig. 1) discovered by an exploratory expedition led by two of us (D. R. and R. J.) in January 1981 shows that some fleeting visits were being made into parts of the region in immediate pre-European times.

Late Pleistocene human occupation east of Tasmania's principal longitudinal watershed has been demonstrated from the Beginners Luck limestone cave site in the Florentine Valley (Fig. 1), an area which was densely forested at the time of European settlement. There, some 20 stone artefacts in a secondary depositional context were initially dated to ~12 kyr BP (ref. 4) but more precisely isolated material has since been radiocarbon dated to $20,650 \pm 1790$ yr BP (ref. 5). Fossils of the extinct megafaunal species *Sthenurus occidentalis* from a nearby site gave an aspartic acid racemization date considerably in excess of this (possibly of the order of 80 kyr), but they indicate radically different environmental conditions in the region some time during the Upper Pleistocene⁶. The discovery of surface stone tools on industrially denuded ridges and moraines on the West Coast Range in 1979⁷, and subsequent unpublished finds by one of us (K.K.) and J. Stockton, further demonstrated the possibility of prehistoric remains existing elsewhere in south-west Tasmania, including the region of the Gordon and Franklin Rivers.

This region offers probably the least archaeological visibility of any in Australia, due to the dense vegetation, rapid peat growth and lack of exposure, but the two rivers run through extensive outcrops of Ordovician limestone where karst landforms are well developed (Fig. 1). Since 1974 a series of pioneering speleological expeditions has documented numerous caves⁸ including in 1977 a large one called Fraser Cave, (F 34), on

the east bank of the Franklin River where original reports noted the existence of an extensive bone deposit⁹. During a later visit in February 1981, Kiernan recognized stone tools and charred bones which indicated this deposit to be of human origin. Accordingly in March 1981 we went up the river to visit the site and conduct a pilot investigation.

Fraser Cave (F 34)

Fraser Cave is a relict outflow stream cave which lies at the head of a resurgence valley 35 m distant from and 10 m above the present level of the Franklin River. It comprises about 200 m of large passages with at least eight surface openings. The largest entrance, which faces north is 12 m wide and up to 3 m high gives easy level access to a spacious chamber (Fig. 2). This chamber is 5 m high, and extends inward for 18 m with a width of between 5 and 12 m. The passages are developed along the strike of the limestone beds which dip to the east at a moderate angle. After having been initiated in phreatic conditions, the cave was further elaborated by the running waters of two small streams which are now intermittent. While some inner sections of Fraser Cave are richly decorated by calcium carbonate speleothems, notably rimstone pools, the downstream areas are dominated by clastic deposits comprising 20–40 cm of moderately rounded fine gravels. These are at present immobile and are subject to manganese encrustation. The gravels are externally derived and consist primarily of Siluro-Devonian metasediments washed from the hillslopes to the east. They are overlain by poorly exposed fine sands which contain at least some stone tools. The sand is in turn overlain by flowstone calcite. This sequence is interpreted as reflecting successive loss of competence by the cave streams. The reduction in clastic load and stream flow permitted precipitation of the flowstone. Piping has since removed some of the sands from beneath the flowstone and a few roof-fall blocks lie scattered on its surface beneath a high level entrance.

Table 1 Radiocarbon assays of various excavation and stratigraphical units

¹⁴ C code no. (ANU)	Excavation unit	Depth below surface (cm)	Stratigraphical unit	Results (yr BP)
2781	A2	3–5	16 top	14,840 ± 930
2782	A8	23–26	13	17,020 ± 310
2783	A10/11	32–39	10	15,670 ± 530
2784	A17/18	79–87	5	Samples too mineralized
2785	A20	94–96	3	19,770 ± 850

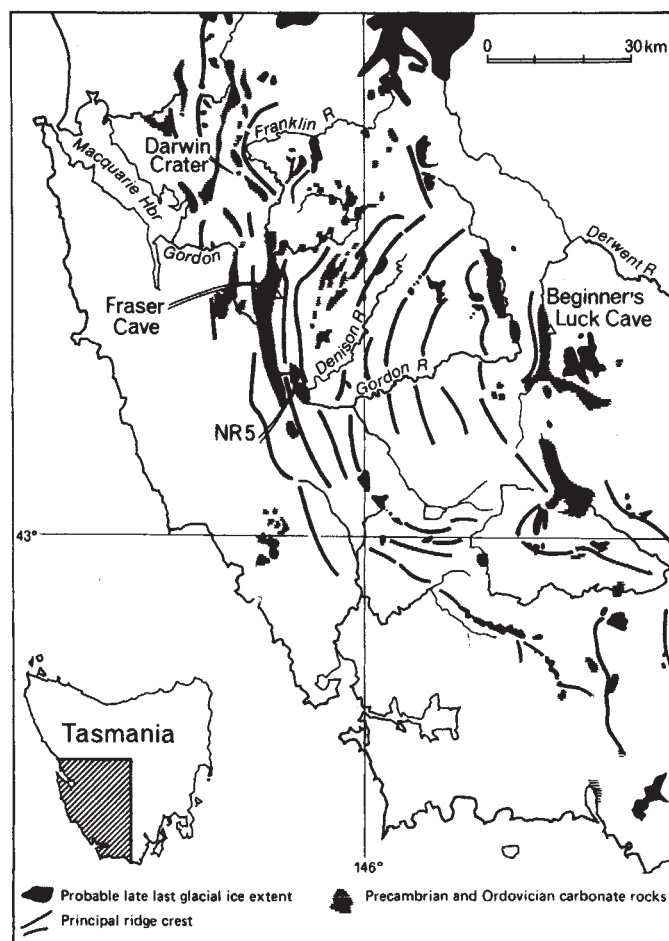


Fig. 1 Map of south-west Tasmania showing main structural features and outcrops of carbonate rocks; also location of glaciers at ~18 kyr ago.

Excavation stratigraphy and chronology

Within the main entrance chamber, and forming a talus at its mouth, is a poorly sorted clastic deposit between 1 and 3 m thick. At a distance of 12 m in from the entrance on the western side of the chamber, this deposit forms a flowstone-capped bank some 0.7 m high. Numerous stone tools, burnt bones and charcoal protrude from the slightly eroded edge of this bank. A small pilot excavation, was undertaken into this face to ascertain the stratigraphical and chronological context of the archaeological remains and to obtain samples for palaeo-environmental analysis. The excavation pit was 1 m wide and reached the rock floor at a maximum depth of 1.30 m. It revealed a complex stratigraphy as shown in Fig. 3. All of the excavated material was wet-sieved using a 3-mm mesh. The charcoal was floated off and stone artefacts and bone material were removed for analysis.

The stratigraphy of these entrance facies may be broadly differentiated into three main complexes. The basal alluvial deposits are similar to those which occur elsewhere in the cave. Gravels are interbedded with lenses of fine sand, while cut and fill structures are common. The lowermost gravel is encrusted by manganese, which indicates the presence of a surface vegetation sufficient to mobilize the mineral before its burial by 10–15 cm of fine laminated sands (unit 2). The most stratigraphically recent gravel (unit 3) has a mean calibre of ~2 cm and contains numerous bone fragments, stone tools and charcoal. A charcoal sample which was obtained from this unit was radiocarbon assayed at $19,770 \pm 850$ yr BP (ANU 2785) (Table 1).

Alluvial sedimentation was subsequently overwhelmed by a major influx (units 6–10) of angular limestone rubble in a loamy



matrix, the rubble stones having a maximum calibre of 10–15 cm. We believe that these units are the product of mechanical weathering in cold climatic conditions. Interspersed between the rubble layers, which are rich in bones and stone tools, are lenses of fine laminated sands (units 5, 7 and 9). These do not contain any cultural debris, and indicate episodic wetness. A radiocarbon assay of charcoal from the upper horizon of this rubble in unit 10 gave a date of $15,670 \pm 530$ yr BP (ANU 2783).

The subsequent black hearth and rubble complex retains much of its general character. The matrix in which the rubble occurs is heavier and richer in clay, which seems to indicate generally damper conditions. Stone tools and bone are extremely abundant and the clay forming the lower boundary of the complex is baked red by prehistoric fires. A radiocarbon assay of $17,020 \pm 310$ yr BP (ANU 2782) was obtained from a hearth unit (unit 13) in the base of this complex. This inversion with the stratigraphically lower sample ANU 2783 cannot presently be explained but further samples will be submitted from the top of this limestone rubble to gain greater precision as to its age.

The top date for the hearth and rubble complex was obtained only 3–5 cm below the surface and gave a result of $14,840 \pm 930$ yr BP (ANU 2781). This dates the uppermost of the horizons which contain tools. Capping this complex is a layer of sterile sand which is 1–3 cm thick, the surface of which is cemented by a crust of calcium carbonate.

The radiocarbon dates indicate the presence of man in Fraser Cave from ~20 kyr BP until ~15 kyr BP. These dates are consistent with the sedimentological evidence which suggests that this occupation corresponded in time to the Last Glacial Maximum. Although it is difficult to separate climatic and anthropogenic effects, the abundance of charcoal in the basal alluvial sequence beneath the entrance facies of Fraser Cave begs the question as to the role that humanly induced fires may have played in destabilizing the slope mantle in the vicinity of the site.

The onset of conditions of extreme cold in western Tasmania is reflected by slope instability at Henty Bridge on the coast-facing margin of the West Coast Range, dated at $23,860 \pm 890$ yr BP (Gak-5594)¹⁰. In the central part of the West Coast Range, 55 km north of Fraser Cave conditions of maximum cold occurred about 5 kyr later. Drifted wood in proglacial silts which immediately predate the arrival of glacial outwash gravels from the Dante Rivulet Valley has been radiocarbon assayed at $18,800 \pm 500$ yr BP (ANU 2533). These dated silts overlie intact flowering specimens of the alpine cushion plant *Donatia novae zelandiae* which imply depression of the treeline to below 230 m (ref. 11). Here a temperate depression of 5.8 °C is indicated by reconstruction of glacial snow-line altitudes. A slightly greater figure of 6.3 °C is indicated 35 km further inland to the east in the headwaters of the Franklin River, where the glacial snowline was depressed to ~1,040 m. In this latter area,

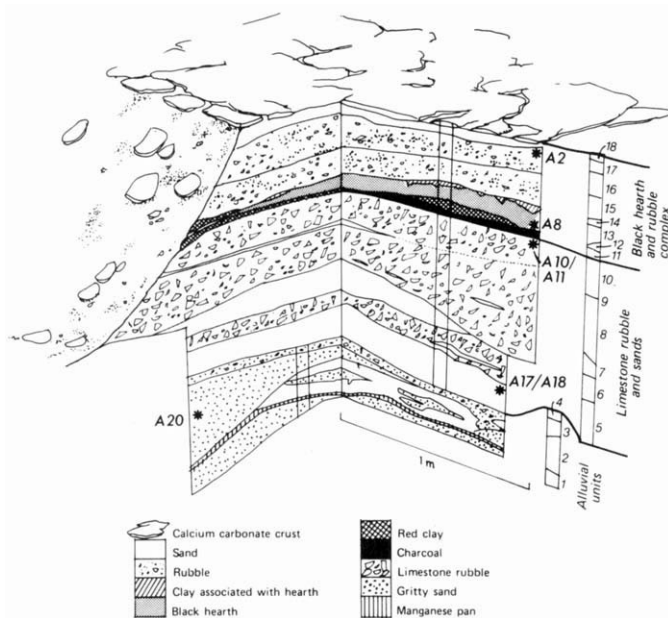


Fig. 3 Stratigraphical section of excavated archaeological deposits at Fraser Cave, showing main units and location of radiocarbon samples.

diffuent ice from the glacier system in the Derwent Valley on the margin of the central plateau ice cap breached the Derwent-Franklin divide and spilled into the Franklin headwaters (Fig. 1). Here it merged with local ice which accumulated in the lee of a snowfence aligned north-south to form a valley glacier which flowed 12 km down the upper Franklin River valley¹¹. Minor cirque and valley glaciers also arose further downstream in the Franklin catchment, notably within the Frenchmans Cap massif 25 km north of Fraser Cave (Fig. 1). Deglaciation appears to have been complete in Tasmania by ~10 kyr BP. The angular limestone rubble in Fraser Cave was produced in cold climatic conditions during this glacial stage and the radiocarbon dates which bracket the rubble and occupation probably broadly delineate the period of maximum cold.

Archaeological remains

Although the total excavation measured ~0.67 m³ in volume, some 75,000 stone flakes and tools were recovered as shown in Table 2. This sample is estimated to be <1% of the total artefact-bearing deposit. Because only ~100 stone artefacts were recovered *in situ* from the only other directly dated Pleistocene sites in Tasmania, namely Beginner's Luck Cave on the Florentine River (Figs 1 and 4) and Cave Bay Cave^{12,13} on present-day Hunter Island (Fig. 4), it can be appreciated that the Fraser Cave assemblage has the potential to transform our present knowledge of Tasmanian late Pleistocene stone technology. The oldest artefact was a single flake recovered from within the basal stratigraphical unit 1, which indicated some occupation probably before 20 kyr BP. However, several hundred flakes and worked tools together with much charcoal were found within the succeeding sands and gravels of unit 3. In almost all the succeeding units, especially the limestone rubble layers, there are numerous stone tools, the only sterile layers being the sand lenses (units 5, 7, 9). In the uppermost complex, there are superimposed hearths with lenses of red-baked clay and abundant charcoal flecks. Occupation ceased suddenly after unit 16 which is dated to ~15 kyr ago. Then the deposit was covered by a thin flow stone, and there is no indication of subsequent human occupation or use of the cave.

The raw materials used for making the stone tools were mostly cobbles of fine-grained siliceous rocks, which could be obtained easily from glacial outwash gravels in the Franklin river-bed nearby. An exception to this lies in a small number of stone tools made from Darwin Glass, which is an impactite associated with a large meteorite crater in the tributary Andrew

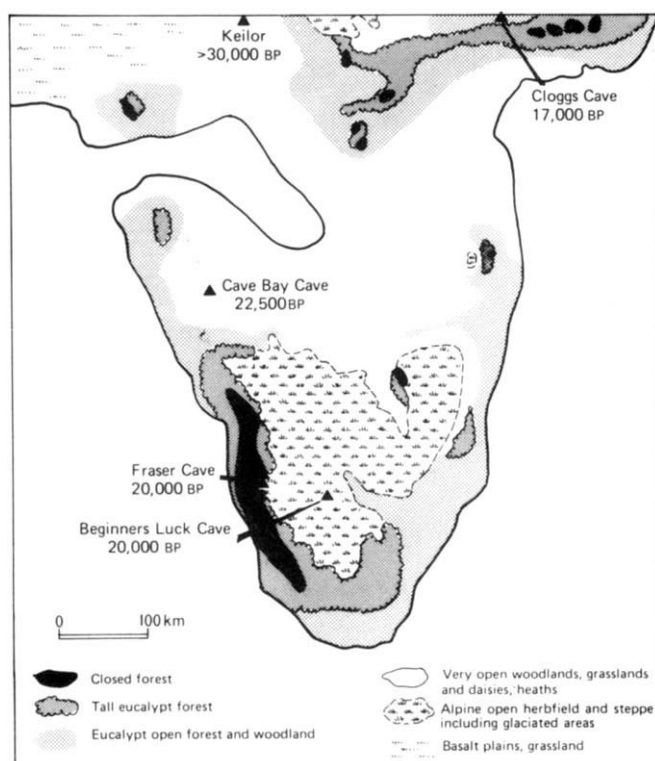


Fig. 4 General reconstruction of major vegetation types and coastline of the Bassian-Tasmanian peninsula at ~18 kyr ago, adapted from ref. 27. Location of prehistoric sites shown with their basal radiocarbon dates.

River Valley 25 km to the north-west (Fig. 1). A westward splash pattern away from the Franklin River has been demonstrated for this material¹⁴, which suggests that these pieces of glass are manuports. Darwin Glass artefacts appear in the sequence from the unit 8 rubble upwards indicating that a considerable mineralogical knowledge of this geologically complex region of south-west Tasmania had been achieved by this time.

Typologically, the tools consist mostly of steep-edge scrapers and domed core-scrapers with steep edges that are often at right angles and show extensive stepped flaking. There are also small round 'thumbnail' scrapers and many retouched flakes. In general this assemblage bears a close resemblance to the tools from the lowest levels of the South Cave, Rocky Cape, on the north coast of Tasmania which have been dated to between 8 and 6 kyr ago¹⁵. The Fraser Cave sequence thus fills a large part of a crucial gap between the base of Rocky Cape and the brief palimpsest of the 20–22-kyr old occupation of Cave Bay Cave^{12,13,16}. From these sites we now have an almost continuous sequence extending from the stone tool technology of the ethnographically recorded Aborigines of the early part of the last century back to just before the Last Glacial Maximum in Tasmania. The Fraser Cave assemblage is also typologically similar to near contemporary industries on the Australian mainland, such as the Lake Mungo assemblage which is dated to ~25 kyr BP (refs 17, 18). All these tools belong to what has been termed the 'Australian core tool and scraper tradition'. It has long been assumed that Tasmanian stone industries were derived from this technological tradition at a time when Tasmania formed part of the single land mass of Greater Australia¹⁹. The Fraser Cave assemblage confirms this. After the post-glacial inundation of Bass Strait, this tradition continued in Tasmania with only slow internal evolutionary changes towards a reduction of average size of tools. On the mainland, however, there were transformational changes associated with the introduction or invention in mid-recent times and later of a variety of small gum-hafted tools such as backed microliths, points and adzes^{20,21}.

Table 2 Occurrence and size of stone flakes, tools and bone fragments

Stratigraphical complex	Stratigraphical units	Vol (m ³)	Stone flakes and tools					Bone fragments	
			Weight		No.		Total	Weight	
			(kg)	(kg m ⁻³)	> 1 cm in length	< 1 cm in length		(kg)	(kg m ⁻³)
Black hearths and rubble	16-11	0.133	10.9	82	900	15,200	16,100	6.0	45
Limestone rubble	10-5	0.318	58.0	182	6,800	51,200	58,000	28.5	90
Alluvial units	4-3	0.089	1.6	18	250	450	700	0.07	8
		0.54	70.5		7,950	66,850	74,800	35.2	

Ochre fragments were found in almost all units above the limestone rubble (unit 6) showing that this pigment was being carried into the cave at this time. Despite an intensive search, no signs of rock art were seen on the walls of the cave.

Bone fragments were found in all of the units which contained stone tools (Table 2), and were absent in culturally sterile layers. From a preliminary analysis based on mandible and maxilla counts, ~90% of the bones are of the large wallaby *Macropus rufogriseus* and about 8% are of the wombat *Vombatus ursinus*. The remaining 2% consist of Tasmanian Devil *Sarcophilus harrisii* and various small mammals. Both Bowdler¹⁶ and Balme²² have independently proposed criteria for distinguishing cave bone accumulations which have resulted from non-human predation such as that by owls and Tasmanian Devils, as opposed to the middens of human hunters. For the latter they suggest an overwhelming preponderance of one or two large game species; a substantial number of bones showing calcination or other evidence of fire; and bones, especially long bones, which have been smashed to obtain the marrow. The fulfillment of these criteria, together with the easy access into and out of Fraser Cave, demonstrate that this bone assemblage is a human midden.

From our preliminary analysis, there do not seem to be any representatives of extinct megafaunal species. This is consistent with the revised analysis of the Beginner's Luck site where the 20-kyr old stone artefacts are associated with bones of modern fauna, mostly *Macropus rufogriseus*, with only one cuboid of a large macropod, cf. *M. titan*⁵. The Fraser Cave assemblage supports the view that by the time of the Last Glacial Maximum, most or all elements of Tasmania's megafauna were already extinct, and that the animals living in these glacial valleys were modern. Such evidence, which contradicts a previously held view^{23,24}, has important implications in the debate concerning a climatic or human causation for the extinction of the 'giant' marsupials. The faunal evidence from the cave also dispels a previous theory that the Last Glacial hunter-gatherers of the Tasmanian peninsula were incapable of systematic hunting of medium-sized land game but were effectively restricted to the width of one band's territory from the coast^{13,25,26}. At the time of its occupation, Fraser Cave would have been ~60 km in a direct line across rugged, snow-bound mountains or 100 km down the most practicable river route from the contemporary coastline (Figs 1 and 4).

The inhabitants of Fraser Cave hunted game with a tight targetting strategy which concentrated on wallabies. These would have been obtained from the open valley slopes which were perhaps cloaked by sub-Antarctic herbfield vegetation. The evidence from the cave suggests a high biomass of these animals having existed in the region during glacial times. The rain-forest may have existed within gallery refugia along the main rivers and possibly in areas closer to the lower coastline including the exposed trough of what is now Macquarie Harbour. This is indicated schematically in Fig. 4 which we derive

from a general reconstruction of the full Last Glacial vegetation by Hope²⁷. We argue that as the climate ameliorated, and conditions became wetter, these forests emerged from their glacial refugia, and despite any effects of man and his fire sticks, re-occupied the higher valley slopes of the Franklin and Gordon rivers. This dense rain-forest habitat was not conducive to hunters, and so Fraser Cave was abandoned, not to be occupied for the next 15 kyr.

Hunters close to southern ice

In addition to the great richness of this site, its significance must be seen in the context of the human colonization of the western rim of the Pacific. The colonization of Australia-New Guinea across the water barriers of Wallacea occurred some time before 40 kyr BP (refs 19, 28). Between at least 50 kyr BP and ~24 kyr BP, the sea was lower than its present level²⁹, but it may not have been low enough to expose the entire Bassian bridge hence the final southward movement of the Australian colonists was stemmed^{25,28}. Then with the onset of the Last Glacial Maximum, the sea dropped below the crucial 60-m isobath, and exposed a dry route to the Tasmanian peninsula. That man immediately seized this opportunity to expand his range is shown by the near contemporary basal occupation of 22.5 kyr BP at Cave Bay Cave^{13,16}, and from possibly slightly earlier than 20 kyr BP at Fraser Cave and Beginner's Luck Cave⁵. In southern Tasmania, these hunters were then the most southerly human beings on Earth. It would not be until 10 kyr BP that equivalent areas in Fuego-Patagonia were to be occupied by man^{30,31}. At 18 kyr ago, the Tasmanians alone were as close to the great Antarctic ice sheet, then only some 1,000 km further to the south³² as some Upper Palaeolithic hunters of Europe, were to the northern ice sheets. The Tasmanian hunters probably lived in tundra environmental conditions similar to those which existed in parts of northern Europe. The specific targetting onto reindeer by the European hunters bears comparison with the similar emphasis on wallabies by the sub-Antarctic palaeo-Tasmanians.

South-west Tasmania therefore offers a fascinating southern analogue for the study of hunters of the Last Glacial Maximum. However, the future of the Franklin-Gordon archaeological sites is threatened by plans now being implemented of the Tasmanian Hydro-Electric Commission to inundate these valleys. This would flood Fraser Cave together with other unstudied sites which have been located within recent months in the wake of the original discoveries.

Our exploratory work in the field in 1981 was supported by the Australian National University, the Tasmanian National Parks and Wildlife Service and the Tasmanian Wilderness Society. We thank our colleagues during the excavation; Barry Blain, Bob Burton, Steve Harris and Greg Middleton; and Hendrick Gout for free helicopter transport of our finds. John Head paid special attention to our dating needs and Winifred Mumford drew the figures.

Received 2 September; accepted 20 October 1982.

1. Jones, R. in *Aboriginal Tribes of Australia* (ed. Tindale N. B.) 319 (University of California Press, San Francisco, 1974).
2. Plomley, N. J. B. (ed.) *Friendly Mission: the Tasmanian Journals and Papers of George Augustus Robinson, 1829-1834* (Tasmanian Historical Research Association, Hobart, 1966).

3. Binks, C. J. *Explorers of Western Tasmania*, 124 (Mary Fisher Bookshops, Launceston, 1980).
4. Goede, A. & Murray, P. *Mankind* **11**, 2 (1977).
5. Murray, P., Goede, A. & Bada, J. L. *Archaeol. Phys. Anthropol. Oceania* **15**, 142 (1980).
6. Murray, P. in *The South West Book* (eds Gee, H. & Fenton, J.) 126 (Australian Conservation Foundation, Melbourne, 1978).
7. Corbett, K. P. *Pap. Proc. R. Soc. Tasmania* **107**, 191 (1980).

8. Middleton, G. J. *Wilderness Caves of the Gordon-Franklin River System* (Centre for Environmental Studies, University of Tasmania, Hobart, 1979).
9. Middleton, G. J. *Sydney Speleol. Soc.* **23**, 51 (1979).
10. Colhoun, E. A. *et al. Quaternary Excursion to Northwest Tasmania* (University of Tasmania, Hobart, 1979).
11. Kiernan, K. thesis, Univ. Tasmania, Hobart (1980).
12. Bowdler, S. *Nature* **252**, 697 (1974).
13. Bowdler, S. in *Sunda and Sahul: Prehistoric Studies in Southeast Asia, Melanesia and Australia* (eds. Allen, J., Golson, J., Jones, R.) 205 (Academic, London, 1977).
14. Fudali, R. F. & Ford, R. J. *Meteoritics* **14**, 283 (1979).
15. Jones, R. thesis, Univ. Sydney (1971).
16. Bowdler, S. thesis, Australian National Univ., Canberra (1979).
17. Bowler, J. M., Jones, R., Allen H. & Thorne, A. G. *Wild Archaeol.* **2**, 39 (1970).
18. Bowler, J. M., Thorne, A. G. & Polach, H. A. *Nature* **240**, 48 (1972).
19. Jones, R. *Nature* **246**, 278 (1973).
20. Mulvaney, D. J. *The Prehistory of Australia*, 210 (Penguin, Melbourne, 1975).
21. Lorblanchet, M. & Jones, R. *Bull. Soc. prehistor. Française* **76**, 463 (1979).
22. Balme, J. *Archaeol. Phys. Anthropol. Oceania* **15**, 81 (1980).
23. Goede, A., Murray, P. & Harmon, R. *Artefact* **3**, 139 (1978).
24. Horton, D. R. *Archaeol. Phys. Anthropol. Oceania* **15**, 86 (1980).
25. Jones, R. in *Sunda and Sahul: Prehistoric Studies in Southeast Asia, Melanesia and Australia* (ed. Allen, J., Golson, J. & Jones, R.) 317 (Academic, London, 1977).
26. Bowdler, S. in *Tasmanian Year Book Vol. 15*, 6 (Hobart, 1981).
27. Hope, G. in *Pleistocene Extinctions; the Search for Causes* (eds Martin, P. S. & Klein, R. G.) (University of Arizona Press, Tucson, 1982).
28. Jones, R. *Ann. Rev. Anthropol.* **8**, 445 (1979).
29. Chappel, J. & Thom, B. G. in *Sunda and Sahul: Prehistoric Studies in Southeast Asia, Melanesia and Australia* (eds Allen, J., Golson, J. & Jones, R.) 275 (Academic, London, 1977).
30. Bird, J. *Geogr. Rev.* **28**, 250 (1938).
31. Borrero, L. A. *Anales del Instituto de la Patagonia* Vol. 8, 81 (Punta Arenas, Chile, 1977).
32. Hays, J. D., Lozano, J., Shackleton, N. & Irving, G. *Geol. Soc. Am. Mem.* **145**, 337 (1976).

Delayed *de novo* methylation in teratocarcinoma suggests additional tissue-specific mechanisms for controlling gene expression

James W. Gautsch & Michael C. Wilson

Committee for the Study of Molecular Genetics, Research Institute of Scripps Clinic, La Jolla, California 92037, USA

Retrovirus infection of embryonal carcinoma cells is blocked at the level of provirus transcription. De novo methylation of the input provirus occurs in embryonal carcinoma cells but not in permissive, differentiated teratocarcinoma. The kinetics of proviral methylation in embryonal carcinoma cells, however, suggest that while methylation may have an important maintenance role in controlling gene expression, additional mechanisms are used by the early embryo to initiate negative gene expression.

THE molecular mechanisms governing the regulation of gene expression during early development of the mammalian embryo are little understood. The introduction of exogenous viral gene sequences into embryonic cells provides a means of examining precisely how this control is initiated and maintained in the uncommitted, pluripotent embryonic cell and how it is subsequently modified after cellular differentiation. As shown by Jaenisch *et al.*¹, the early embryonic cells of the mouse are refractory to infection with Moloney murine leukaemia virus (M-MuLV). Similarly, a dramatic block to virus replication is observed after infection of embryonal carcinoma (EC) cells with M-MuLV². As the EC cells are an *in vitro* model of the undifferentiated preimplantation mouse embryo cell (for review see ref. 3), the mechanisms governing this control of viral gene

expression are probably the same. Teich *et al.*⁴ have shown that although progeny virus cannot be detected, M-MuLV is successfully adsorbed and viral DNA is introduced into the EC cell. Fusion of infected EC cells with differentiated cells⁵ or karyoplasts, but not cytoplasts⁶, results in virus production, suggesting that the EC cell either lacks a function necessary for virus production or contains an inhibitor specifically limited to the EC cell nucleus. Speers *et al.*⁷ showed that infected and cloned EC cell lines that were not producing virus could be made to do so after differentiation and exposure to 5-bromodeoxyuridine (BUDR). Input M-MuLV was recovered from two of four infected and cloned EC cell lines several months after infection. This provides evidence that proviral DNA made from input M-MuLV RNA persists indefinitely in

Table 1 Transcription rate of M-MuLV homologous RNA in EC and differentiated teratocarcinoma cells infected with M-MuLV

Time of pulse label (min)	EC ⁻ MuLV	EC ⁺ MuLV	Diff. EC ⁺ MuLV	EC ⁻ MuLV Diff. EC ⁺ MuLV	EC ⁺ MuLV Diff. EC ⁺ MuLV
c.p.m.	125	86	33,752		
10 % Hybrid†	0.0018	0.0012	0.312	0.0058	0.0038
c.p.m.	139	106	74,242		
30 % Hybrid	0.0005	0.0004	0.226	0.0022	0.0018

EC⁻MuLV were uninfected EC-B cells; EC⁺MuLV was the non-productively infected EC-BMBcl₂ cell line; differentiated (Diff.) EC⁺MuLV was a M-MuLV148-infected differentiated teratocarcinoma cell line, SSCM. Cells in 150-mm diameter tissue culture plates were labelled with [³H]uridine at 160 µCi per ml of medium for 10 or 30 min at 37 °C. RNA transcription was stopped after 10 and 30 min with ice-cold phosphate-buffered saline (PBS). Cells were lysed with 0.2% Nonidet P-40 (NP40) and pelleted nuclei were resuspended and lysed in 0.5 M NaCl with DNase at 37 °C for 1 min and hot phenol-extracted. RNA was precipitated sequentially with ethanol, LiCl₂ and ethanol. Labelled nuclear RNA was hybridized for 40 h at 67 °C to an excess of pMLV-1 DNA bound to nitrocellulose filters. The pulse-labelled M-MuLV-specific RNA was measured after treatment with pancreatic and T₁ RNase as described previously³⁰.

* c.p.m. hybridized after subtraction of background blank filters.

† % Of total labelled RNA in hybrid to pMLV-1.

the EC cells either as an episome or integrated into the EC cell genome. These and biochemical studies by Teich *et al.*⁴ and D'Auriol *et al.*⁸ did not unequivocally resolve the question of whether input viral DNA was or was not integrated into the genome of EC cells; however, Stewart *et al.*⁹ recently demonstrated the presence of multiple proviruses after infection of F9 cells with M-MuLV.

We show here that infection of undifferentiated PCC4aza1 EC cells with M-MuLV also results in integration of proviral DNA into the genome of the EC cell, but proviral DNA is neither transcribed nor is it infectious as a transfection donor. We find that proviral DNA in infected non-producer EC cell lines is methylated at sites that are not methylated after infection of differentiated cells which support active viral gene expression. By following the kinetics of methylation after infection of EC cells we find, however, that provirus is not methylated until after 8 days post-infection and thereafter it occurs slowly. Little or no viral gene transcription is detected after infection prior to methylation. These results, therefore, define a control of gene expression that may involve *de novo* methylation as well as additional mechanisms.

Infecting viral DNA integrates into embryonal carcinoma cells

To determine the structure of the exogenous M-MuLV provirus within the nucleus of the EC cell, we first required a method that would distinguish between the M-MuLV DNA and multiple retroviral sequences endogenous to the mouse genome^{10,11}. Examination of the nucleotide sequence of M-MuLV¹² showed that *Bam*HI cleavage of the linear proviral genome would result in two internal fragments of 3,001 and 297 bases (see Fig. 1B). Similarly, *Bam*HI digestion of the circular form of the M-MuLV genome, found before integration, would generate a fragment 5.0 or 5.6 kilobases (kb) long, dependent on the presence of one or two long terminal repeat (LTR) junctions. Figures 1A and 3A show the results obtained with DNA preparations of uninfected and infected EC and differentiated teratocarcinoma cell line DNAs hybridized to cloned M-MuLV DNA¹³. Hybridization to the M-MuLV-specific 3.0-kb *Bam*HI fragment was observed only in the DNA of M-MuLV-infected cells. Inspection of the *Bam*HI digests in Figs 1A and 3A, B showed no evidence of new fragments of 5.0 or 5.6 kb long, indicative of circular, unintegrated viral structures, in either EC or differentiated cells infected with M-MuLV. Similarly, when fractionated without restriction enzyme cleavage, the same DNA preparations showed no evidence of viral DNA at 8.8 or 5–6 kb, the migration position found in agarose gels for unintegrated linear and circular supercoiled proviral DNA, respectively^{8,14}. We concluded that the M-MuLV provirus is integrated in the infected but non-producer EC cell lines. Restriction of virion production in EC cells must, therefore, be maintained by mechanisms other than coordinate integration of provirus and virion production during cellular differentiation.

Based on the intensity of hybridization of the M-MuLV-specific 3-kb band, we were able to compare the copy number of proviral sequences in these cells with the numbers of provirus in cloned lines of permissive, differentiated teratocarcinoma cells which produce M-MuLV. The other bands of hybridization found in the uninfected EC cells DNA after *Bam*HI digestion represent fragments of endogenous retroviral sequences present in multiple copies within the 129 mouse genome^{10,11}. The hybridization of these sequences therefore provides an internal control for the measurement of the number of integrated M-MuLV proviral genomes revealed as the 3-kb *Bam*HI band. Figure 1A gives the relative intensities of the M-MuLV-specific 3-kb fragment compared with an endogenous 2-kb fragment, and reveals that the two productively infected, differentiated cell lines, SSCMG5 and SSCMG8, contain 3 and 10 times the number of M-MuLV genomes present in the non-producer infected EC cell line, BMBcl₂.

In addition to BMBcl₂, DNA from three other subclones (BMA, C and D) of infected EC cells had been previously tested for recovery of virus after differentiation and BUdR induction⁷. Two yielded infectious M-MuLV whereas the other two did not. Of these four cell lines, three showed the *Bam*HI M-MuLV-specific band (Fig. 3A). One of the cell lines, BMD, shows evidence of integrated M-MuLV provirus by hybridization although we were unable to recover infectious virus from the cells after differentiation and induction⁷. DNA from BMA did not show a M-MuLV-specific *Bam*HI band, nor could virus be recovered after differentiation and induction. It is possible that BMD contains an integrated M-MuLV provirus that is defective for replication. Alternatively, the M-MuLV provirus may be integrated into the EC cell genome in a region more stringently controlled than the proviral positions in the BMB and BMC EC cell clones from which virus could be recovered.

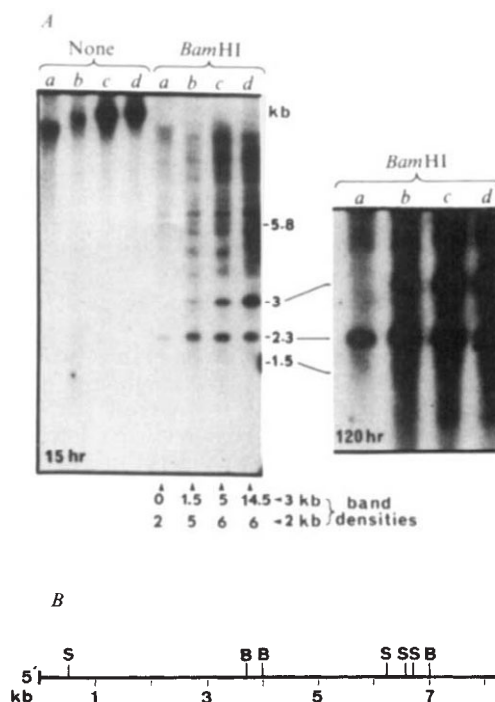


Fig. 1 A, analysis of integrated or non-integrated state of M-MuLV provirus in EC and differentiated cells. Lanes a show DNA isolated from uninfected EC-B cells, a subclone of PCC4aza1³². PCC4aza1 is an EC cell line established from teratocarcinoma OTT6050A which arose in a 129/Sv mouse³³. Lanes b show DNA from the BMBcl₂ subclone of M-MuLV 148-infected EC-B cells. Lanes c and d are DNA from subclones SSCMG5 and SSCMG8, respectively—two cell lines isolated from productively-infected differentiated teratocarcinoma cell lines. SSCMG5 and SSCMG8 were derived by cloning single cells after infection of SSC cells with M-MuLV 148. SSC is a cell line isolated by Speers from an *in vitro* dimethylacetamide-treated EC-B culture³⁴. High-molecular weight chromosomal DNA was isolated from nuclei of monolayer cultures by lysis with 0.5% SDS, proteinase K and RNase A digestions, phenol extractions and ethanol precipitations. Lanes a contained ~5 µg DNA and lanes b–d contained ~15 µg DNA each. The DNA was size-separated in 0.6% agarose and transferred to nitrocellulose. The nitrocellulose-bound DNA was then hybridized to nick-translated ³²P-labelled pMLV-1 DNA¹³. *Bam*HI-digested pMLV-1 fragments run on the same gel were used as size markers. A section of the *Bam*HI-cut DNAs is shown after a 120-h exposure (8 times longer) to determine clearly the lack of a M-MuLV-specific 3-kb band in lane a. The relative densities of the 2- and 3-kb bands were measured with a soft laser scanning densitometer (Zeineh) and the values are given under the respective lanes. B, restriction map of M-MuLV provirus based on DNA sequence¹². S = *Sma*I/*Xba*I, B = *Bam*HI cleavage sites. The LTRs on each end are identical and consist of 592 bp. The provirus is 8,854 bp long.

Provirus is not transcribed in EC cells

To establish whether the restricted expression of M-MuLV in the non-producer EC cell is due to the lack of proviral mRNA transcripts, polyadenylated RNA was electrophoretically fractionated on denaturing agarose gels and analysed by RNA blot hybridization¹⁵. The steady-state level of viral RNA was determined by hybridization with the ³²P-labelled M-MuLV probe, pMLV-1. In these experiments, controls for the mass of RNA applied to the gel and for RNA degradation were provided by the cellular mRNA probes, pCHOB and pCHOC¹⁶.

Figure 2 shows the results of two experiments in which we examined the viral RNAs of infected and uninfected differentiated and undifferentiated cells. Viral homologous mRNAs could be readily visualized in the poly(A)⁺ RNA population of differentiated cells producing infectious viral particles, represented by the Scripps 60A cell line¹⁷ (Fig. 2A), and in the infected differentiated EC-derived cell line, SSC (Fig. 2B). In the M-MuLV-infected SSC cell, the viral RNA consists of several poly(A)⁺ transcripts which vary in size from 8.0 to 3.5 kb (refs 18, 19). No viral RNA was, however, detected in the undifferentiated infected EC cell, even after long exposure (Fig. 2A, lane b and 2B, lane e). The hybridization of the control probes, pCHOB and pCHOC, indicates that the RNA isolated from the EC cells is intact and present in an amount approximately equal to that isolated from the productively infected cells. We therefore conclude that the expression of integrated M-MuLV provirus in the non-producer EC cell as stable mRNA is below the level detectable by this technique, and is probably less than 10 copies per cell.

To determine whether the expression of the MuLV proviral sequence is governed at the level of transcription or by potential post-transcriptional mechanisms, we pulse-labelled nascent nuclear RNA transcripts of undifferentiated infected non-producer EC cells and the differentiated cell line SSC, infected with M-MuLV. The total nuclear RNA of cells labelled for either 10 or 30 min was then assayed for labelled M-MuLV transcripts by hybridization with excess pMLV DNA. Table 1 shows that the level of viral RNA that hybridizes to pMLV-1 DNA in infected producer cell nuclei is 400–700-fold higher than that in infected non-producer EC cells. The level of viral RNA in the latter cells was equal to the background level of endogenous proviruses in uninfected cells of 129 mice¹⁸. The results indicate, therefore, that the provirus is not transcribed in BMBcl₂ cells and that this is the level of the EC cell block to M-MuLV infection.

Provirus in EC cells is not infectious

The mechanism(s) of the EC cell transcriptional block of M-MuLV proviral DNA is of interest because it defines a fundamental difference between EC cells and almost all other mouse cells in the manner in which genes are expressed. Does the EC cell fail to provide a factor necessary for transcription or is the proviral DNA rendered transcriptionally silent by a *cis* or *trans* inhibition mechanism operating in the EC cell nucleus? One means of determining whether provirus is modified in some manner is to test its ability to infect permissive cells productively in a transfection assay. DNA from M-MuLV-infected differentiated producer cells and the undifferentiated non-producer EC cell line, BMBcl₂, was isolated and transfected into NIH 3T3 cells according to the CaPO₄ precipitate method of Graham and van der Eb²⁰. As shown in Table 2, infectious M-MuLV was recovered from NIH 3T3 cells transfected with infected differentiated cell DNA isolated from SSCMG5 and SSCMG8 cells which contained at least 3 and 10 copies of M-MuLV provirus, respectively, and from M-MuLV-infected NIH 3T3 cells. No infectious virus was detected in cultures of NIH 3T3 cells transfected with DNA from BMBcl₂ or uninfected EC cells. At 20 µg DNA per plate, the virus producer cell DNAs gave 100% positive transfection. SSCMG5 DNA at

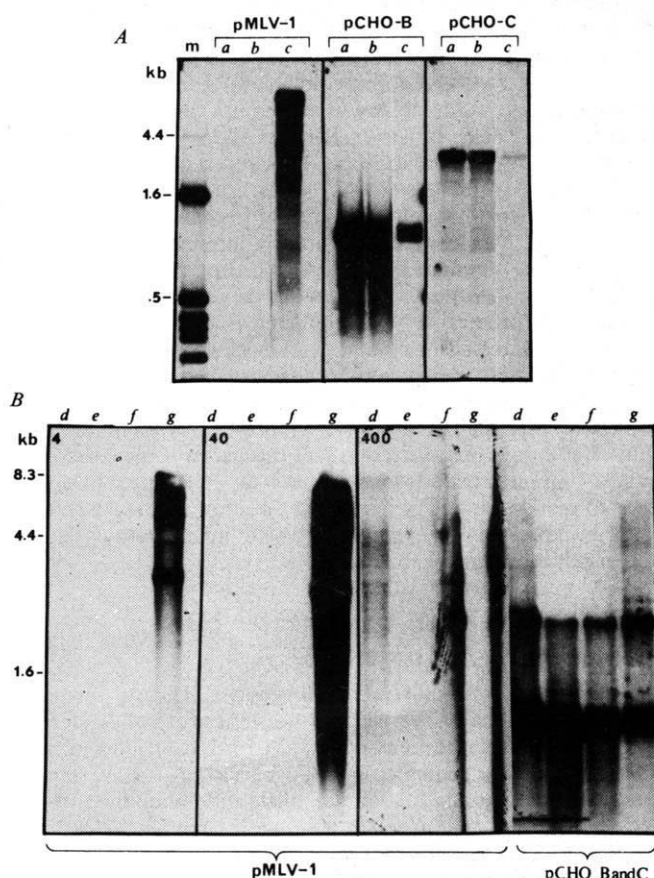


Fig. 2 Analysis of poly(A)⁺-containing RNA from infected and uninfected EC and differentiated cell lines. Poly(A)⁺-containing RNA was isolated from the cell cytoplasm as described previously²⁰. Monolayer cultures in 150-mm diameter tissue culture plates were washed with cold PBS and lysed with 0.2% NP40. The nuclei were pelleted from the lysate, the supernatant was extracted with phenol and chloroform, and the RNA precipitated with ethanol. Poly(A)⁺-containing RNA was selected on oligo(dT) cellulose columns and ethanol precipitated. We experienced difficulties in obtaining undegraded RNA from differentiated teratocarcinoma cells by the above methods but found that SDS lysis of the monolayer cultures in the presence of proteinase K eliminated RNase activity and allowed isolation of intact poly(A)⁺ RNA in these cases. Poly(A)⁺ RNA (5–10 µg) dissolved in MOPS buffer containing 50% formamide and 2.2 M formaldehyde was electrophoresed through 1.5% (A) or 1% (B) agarose, in the presence of 1.1 M formaldehyde at 150 V for 4–5 h. The RNA was transferred to nitrocellulose as for DNA and hybridized to nick-translated ³²P-labelled DNA probes²⁵. To detect MuLV mRNA, pMLV-1 was used as probe. Duplicate RNA samples were also hybridized to labelled pCHOB and pCHOC probes as a measure of the relative quantity of RNA in each sample. These probes were derived from cDNAs of multiple-copy 'housekeeping' mRNA species of Chinese hamster ovary cells¹⁶ that cross-hybridize with conserved mRNA sequences in cells of different species (M.C.W., unpublished observations). In A: a, 6 µg of EC-B; b, 5 µg of EC-BMBcl₂; and c, 3 µg of 60A cell cytoplasmic RNAs were electrophoresed, transferred to nitrocellulose and hybridized with the designated probes. Although a large quantity of mRNA in EC cells hybridizes to CHOB, no viral mRNA is detectable in these cells, whereas species of viral mRNA are seen in 60A cells, a MuLV-producing lymphocyte cell line. The autoradiogram shown is overexposed (3-day) to show no viral RNA present in these cells at this detection level relative to other cellular mRNA. B shows three different time exposures of an autoradiogram of RNA from infected and uninfected cell lines hybridized to pMLV-1 and the 'housekeeping' cDNA probes pCHOB and pCHOC. e, EC-BMBcl₂; f, uninfected, differentiated teratocarcinoma SSC; and g, SSC cells productively infected with M-MuLV-148. The wide range of quantities of viral-specific RNA in these hybridizations was best represented as autoradiograms exposed to film for various times. At 400 h exposure, background hybridization to endogenous provirus mRNA was just visible in EC-BMBcl₂ (e) and SSC (f) cells, where a 4-h exposure was sufficient to detect viral RNA in infected producer SSC cells (g). Lane g was excised before exposure for 400 h to eliminate masking of adjacent lanes. Based on the range of band intensities in B, there was ~10⁴-fold more hybridization to viral RNA in infected SSC than in infected or uninfected EC cells or uninfected differentiated cells. The rightmost section of B shows 400-h exposure of the hybridization to pCHO-B and -C probes. Nicking of RNA in lane e is seen as a smaller upper pCHOC band and 'smearing' below the pCHOB band. This is reflected in the hybridizations to pMLV-1 and monitors intact mRNA. Sizes (kb) determined from digested pBR322 DNA fragments separated on the same gels.

Table 2 Transfection of NIH 3T3 cells with DNA from infected non-producer EC cells and infected differentiated producer cells

DNA from donor cells	μg DNA added per plate			
	1	5	10	20
		XC ⁺ /Total		
EC-A	0/2	0/3	0/2	0/4
BMBcl ₂	0/1	0/6	0/6	0/8
SSCG5	0/2	2/6	4/6	5/5
SSCG8	1/2	2/6	5/5	3/3
NIH 3T3-148	0/2	3/3	4/5	4/4
EC-A (5 days p.i.)			1/3	

Chromosomal DNA was isolated as described in Fig. 1 legend from the cell lines indicated. EC-A (5 days p.i.) were EC cells 5 days after infection with M-MuLV; the M-MuLV provirus in these cells was not yet hypermethylated (see Fig. 4) but no virus was being produced by the culture. DNA was sheared by passage through a 21-gauge needle (first five) or cleaved with *EcoRI* (last one) which does not cut M-MuLV provirus. The fragmented DNA was added to a monolayer of 3×10^5 NIH 3T3 cells growing in 35-mm diameter tissue culture plates. The DNA was added with a CaPO₄ precipitate formed in the DNA solution according to established procedures²⁰. The cells were washed with medium 4 h later and 2×10^5 SC-1 cells were added to each plate. The plates were incubated in the presence of $4 \mu\text{g ml}^{-1}$ polybrene for 2 days. Aliquots were co-cultivated in separate dishes with XC cells in medium containing μM dexamethasone³¹ and 3 days later the presence (XC⁺) or absence of XC syncytial plaques was observed. The original cultures were maintained and tested for XC plaques periodically; if XC-negative 2 weeks after adding DNA, the transfection was considered negative. Each value shown, therefore, represents a plate of NIH 3T3 cells to which DNA had been added and which was subsequently determined to be positive or negative for production of M-MuLV.

concentrations between 1 and 20 μg per plate gave 11 virus-positive cultures out of 19 cultures exposed to the DNA. The DNA of the infected non-producer EC cell line, BMBcl₂, which contains at least one provirus copy, should have produced one-third as many transfection-positive cultures, or at least three, but none was found; this strongly suggests that the provirus in BMBcl₂ is inactivated in some way. This inactivation, however, must be reversible because after differentiation, infectious M-MuLV can be recovered from differentiated BMBcl₂ cells⁷. These results agree with others which have recently shown that M-MuLV proviral DNA in mouse^{21,22} and F9 (ref. 9) cells is non-infectious in transfection assays if the virus is not expressed at the time of DNA isolation and the provirus is methylated. It is probable that the provirus in BMBcl₂ cells is not infectious for similar reasons.

M-MuLV provirus in subclones of non-producer EC cells is methylated

The proviral DNA in infected differentiated and non-producer EC cells was examined for methylation by testing for resistance to restriction enzyme cleavage using enzymes that are sensitive to m⁵CpG methylation. From the sequence of M-MuLV¹², we know that *HpaII*, which recognizes and cleaves CCGG but not CmCGG, has 34 recognition sites in M-MuLV proviral DNA, resulting in 33 small internal fragments, many of which are about the same size and would not be easily resolved. There are, however, five *SmaI* recognition sequences (CCCGGG) in M-MuLV. Cleavage by *SmaI* is also precluded by methylation of the internal C residue. The enzyme *XmaI* shares the same recognition sequence as *SmaI*; however, its activity is not sensitive to methylation²³. As predicted from the DNA sequence of

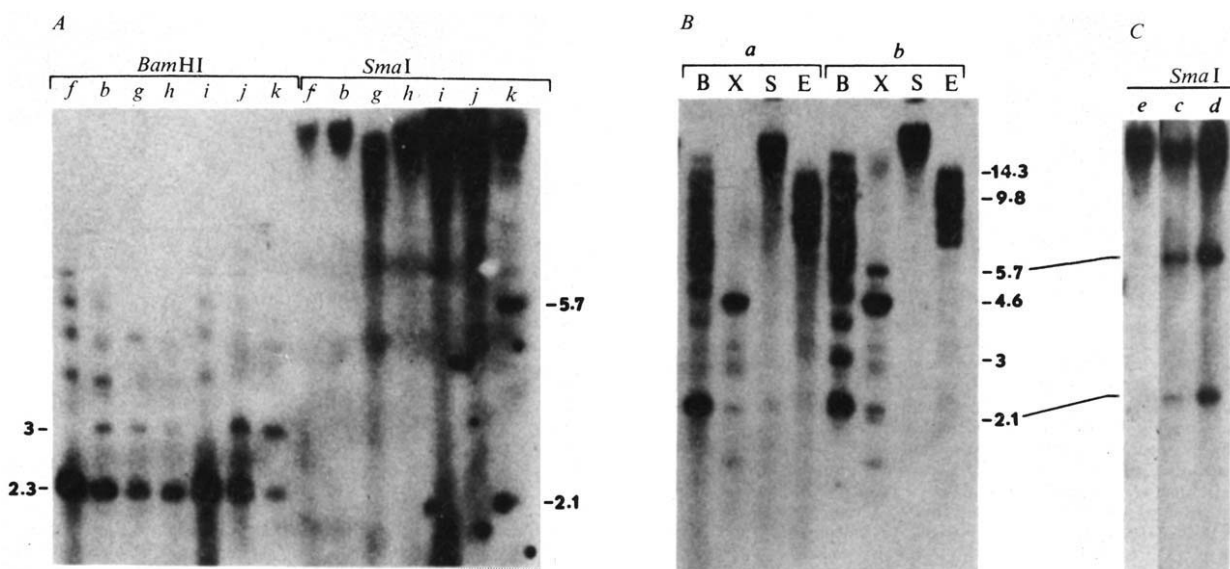


Fig. 3 A, subclones of PCC4azal, EC-A and EC-B, were infected with M-MuLV-148 at a MOI of 2 and after several weeks, single cell clones were isolated from these cultures and analysed for the presence of M-MuLV provirus. Each lane contained 5–10 μg nuclear DNA cleaved to completion with *BamHI* or *SmaI*. DNA was separated on a 0.8% agarose gel and hybridized to pMLV-1 as described in Fig. 1 legend. The lanes contain DNA from EC cell clones BMA(f), BMBcl₂(b), BMC(g), BMD(h), AMG5(i), AMG12(j) and SSD148(k), an infected differentiated teratocarcinoma cell line, SSD, from Speers³⁴. *BamHI* cleavage of the DNA revealed the presence of provirus if the M-MuLV-specific 3.0-kb band was seen. None of the infected EC cell cultures, cloned or uncloned, produced M-MuLV. The *BamHI* 3.0-kb band is evident in all lanes except f and i. These two subclones do not contain M-MuLV provirus. In general, we found evidence of M-MuLV provirus in 75% of the subclones of EC cells after infection at a MOI of 2. Although DNAs in lanes b, g, h, and j contain M-MuLV provirus, they do not show the 5.7- and 2.1-kb M-MuLV-specific *SmaI* bands, indicating that the provirus is methylated at CCCGGG sites. B, a, b, DNA from uninfected EC cells and BMBcl₂ cells, respectively, was digested with restriction enzymes *BamHI* (B), *XmaI* (X), *SmaI* (S) and *EcoRI* (E). Aliquots of the reaction mixtures were mixed with λ phage DNA to monitor the reaction end point. The enzyme-cleaved DNA was separated by size on a 0.8% agarose gel and hybridized to pMLV-1 as described above. Band sizes were determined by plotting migration distance in log kb using *HindIII*- and *EcoRI*-digested λ C₁₈₅₇ SAM7 DNA as standards. M-MuLV-specific bands are indicated as follows: 3 kb in lane b, B; 2.1 and 5.7 kb in lane b, X; and most probably 14.3 kb in lane b, E. These bands are all absent from the uninfected EC cell DNA in lanes a. The 2.1- and 5.7-kb bands are also absent from lane b, S, indicating that the M-MuLV provirus is methylated at CCCGGG sites. C, DNA from uninfected, differentiated teratocarcinoma cell line SSC (e), and infected SSC clones, SSCMG5 (c) and SSCMG8 (d), were cleaved with *SmaI* to completion, separated and hybridized to pMLV-1 as described for B. The 5.7- and 2.1-kb M-MuLV-specific *SmaI* proviral fragments are evident in the infected cell DNAs, indicating a lack of methylation in these sites in DNA from differentiated cells permissive for retrovirus infection. Without exception, provirus from all cell lines tested was cleaved by *SmaI* if the cells were producing virus, and provirus from subclones of infected non-producer EC cells was resistant to *SmaI* cleavage.

M-MuLV, cleavage of the provirus by these enzymes would generate four internal fragments of 5,720, 2,059, 329 and 156 base pairs (bp). Figure 3B shows that pMLV-1 DNA hybridization to *Sma*I-cleaved, infected, non-producer EC cells and uninfected EC cells did not result in bands at 5.7 and 2.1 kb. *Xma*I digestion of the same DNAs showed M-MuLV-specific 5.7- and 2.1-kb bands in BMBel₂ DNA which were absent from the uninfected EC cell DNA. *Xma*I digestion also produced at least 12 additional bands visible in both infected and uninfected EC cell DNA, notably a 4.6-kb band that probably resulted from numerous endogenous proviruses having identical CCCGGG sites. None of these bands in *Xma*I-cleaved DNA was visible in the *Sma*I-digested DNA, indicating that endogenous proviruses, including M-MuLV in this case, are methylated at these sites and almost all pMLV-1 hybridization occurs in uncleaved DNA at the top of the gel. Figure 3A shows an additional three subclones of infected EC cells which contain the 3-kb *Bam*HI fragment of M-MuLV provirus but which are resistant to *Sma*I cleavage. Figure 3A, C shows that the 5.7- and 2.1-kb M-MuLV-specific bands are seen in *Sma*I-cleaved DNA from SSD148, SSCMG5 and SSCMG8, cells which are active in M-MuLV production. Therefore, in cells that are producing M-MuLV the provirus is not methylated at CCCGGG sites, whereas it is methylated in infected non-producer EC cells. In agreement with the recent findings of Stewart *et al.*⁹, we conclude that *de novo* methylation of M-MuLV provirus occurs in EC cells but not in differentiated cells permissive to productive retrovirus infection.

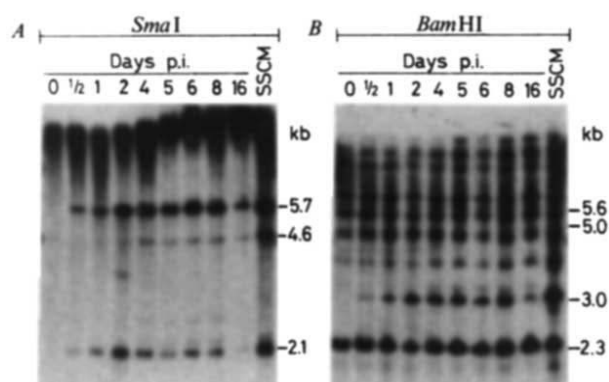


Fig. 4 Measurement of the transient lack of methylation of the M-MuLV provirus early after infection of EC cells. EC cells were infected with M-MuLV (2 PFU per cell). After the indicated intervals post-infection (days p.i.) the nuclear DNA was isolated and aliquots (5 μ g) were digested separately to completion with *Sma*I and *Bam*HI. The DNA fragments were fractionated on a 0.75% agarose gel. Also included were samples of DNA of the productively infected differentiated cell line, SSCM. After blotting to nitrocellulose and prehybridization, the DNA was hybridized at 67 °C in 0.3 M NaCl to pMLV-1, ³²P-labelled to a specific activity of 1.1×10^9 c.p.m. μ g⁻¹ by nick-translation³⁵. After washing, the filters were exposed for 5 days to pre-flashed X-ray film with an intensifying screen. A, the 5.7-kb and 2.1-kb bands are generated by internal *Sma*I cleavage sites specific to the M-MuLV genome (see Figs 1b, 3) and are not found in the uninfected EC cell (lane 0). The 4.6-kb and other minor bands represent endogenous retroviral DNA fragments cleaved by *Sma*I. The intensity of the 5.7- and 2.1-kb M-MuLV provirus bands relative to that of the 4.6-kb endogenous provirus band, is a measure of the degree of M-MuLV provirus methylation. By this measure, a portion of the population of M-MuLV proviruses is methylated at 16 days relative to earlier time points. Thus, methylation of M-MuLV provirus seems to commence after 8 days post-infection but is not complete by 16 days. B, the 3-kb M-MuLV provirus-specific band is not found in the uninfected EC cell DNA (lane 0) but is seen in all other lanes containing DNA from infected EC and differentiated cells. The 3-kb band is a measure of provirus (integrated) as opposed to unintegrated viral DNA because: (1) no bands corresponding to the various linear or circular DNAs were detected in uncut DNA hybridized to pMLV-1 (not shown); and (2) no additional bands were present in the infected versus uninfected DNA at 5.0 and 5.6 kb or 1.8 and 3.6 kb, which would result from *Bam*HI cleavage of circular or linear forms of unintegrated viral DNA, respectively.

Methylation may be the effect, not the cause, of negative proviral gene expression

Because the above demonstration of proviral methylation in EC cells was done using cloned cell lines, we have not determined whether the lack of gene expression is due to DNA methylation or rather the cause of the methylation. We therefore examined the kinetics of *Sma*I site methylation in provirus at various times after virus infection. EC cells were infected with M-MuLV at a multiplicity of infection (MOI) of 2 plaque-forming units (PFU) per cell and chromosomal DNA was isolated at various intervals post-infection and examined for methylation in provirus. Surprisingly, Figure 4A shows that the *Sma*I sites generating the characteristic 5.7- and 2.1-kb fragments of the M-MuLV provirus are sensitive to cleavage for 16 days post-infection. That the *Sma*I fragments are generated from intact, integrated M-MuLV provirus is demonstrated by a parallel *Bam*HI digest of the same DNA samples. Figure 4B shows that the 3-kb *Bam*HI fragment characteristic of the M-MuLV genome is present in all the cultures infected with M-MuLV. Comparison of the intensity of this M-MuLV-specific fragment with that of the 2.3-kb fragment generated from multiple endogenous retroviral sequences indicates a similar number of M-MuLV genomes as present within the chromosomal DNA of the non-producer EC cell lines shown in Figs 1, 3. As both the *Bam*HI and *Sma*I digests shown in Fig. 4 used an identical amount of DNA, were hybridized in parallel with the same labelled probe and exposed for the same length of time, the relative intensity of the M-MuLV-specific hybridization provides a measure of the number of proviral genomic fragments generated by each digest. Comparison of

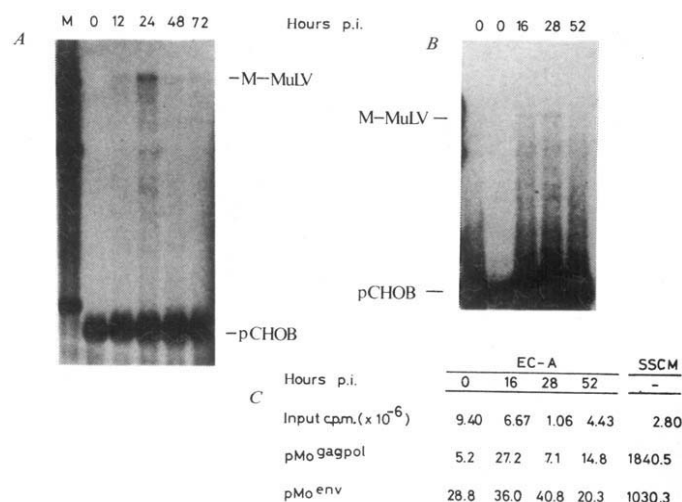


Fig. 5 Non-methylated M-MuLV provirus expression was measured early after infection. After the indicated time post-infection with M-MuLV (at 2 PFU per cell), cultures of EC cells were collected, and nuclei and cytoplasmic fractions obtained after NP40 lysis. Poly(A)-containing RNA was isolated from the phenol-extracted cytoplasmic RNA by oligo(dT) chromatography and 5 μ g of RNA were fractionated on either 1% (A) or 1.25% (B) agarose-formaldehyde gels as described in Fig. 2 legend. After blotting to nitrocellulose and prehybridization, the bound RNA was hybridized to a mixed probe consisting of plasmid clones pCHOB and pMLV-1 labelled with ³²P by nick-translation. The filters were then exposed to preflashed X-ray film with an intensifying screen. The exposures shown are: A, 4 days and B, 20 h. C, transcription rate analysis; $\sim 2 \times 10^7$ nuclei from the same cells whose cytoplasmic poly(A)⁺ RNA was analysed by blot hybridization (shown in B) and the nuclei of M-MuLV-infected SSC cells were incubated in the presence of 1 mM each of ATP, CTP and GTP and 200 μ Ci ³²P-UTP in a 50- μ l reaction volume. After incubation at 25 °C for 45 min, the labelled nuclear RNA was extracted essentially as described by McKnight and Palmiter³⁶ and hybridized to 12 μ g of the plasmid DNAs, pMo^{gag/pol} and pMo^{env} immobilized on nitrocellulose filters. These cloned sequences are subclones of the parental pMLV-1 and represent 5.3 and 2.7 kb of the *gag/pol* and *env* regions of the MuLV genome, respectively¹⁸. The RNase-resistant hybrids were counted and expressed as c.p.m. per c.p.m. $\times 10^6$ of input radioactivity.

Fig. 4A and B indicates that most, if not all, of the M-MuLV proviral genomes are not methylated at the *Sma*I sites during the first 8 days after infection.

Importantly, as found for the infected non-producer cell lines including BMBcl₂, M-MuLV-specific *Bam*HI fragments of 5.0 or 5.6 kb were not detected in DNA preparations, which would be indicative of circular unintegrated viral DNA. Similar examination by Southern blotting of undigested samples of these DNA preparations has also shown no evidence of unintegrated viral DNA (data not shown). These DNA preparations were isolated by spooling the high molecular weight DNA after ethanol precipitation. The presence of infectious unintegrated viral DNA in EC cells after recent infection by M-MuLV has been adequately reported elsewhere^{8,9}. By only isolating high-molecular weight DNA, we were able to examine exclusively the integrated copies of M-MuLV unobscured by the multiple unintegrated forms present before integration. We therefore conclude that shortly after infection of EC cells, M-MuLV DNA integrates but is not methylated at CCCGGG sequences. As the relative intensities of the M-MuLV-specific *Sma*I-generated fragments of the 16-day infected culture seemed to diminish, we suggest that methylation of the newly integrated provirus begins to occur between 8 and 16 days post-infection.

Because unmethylated copies of the provirus were present within the genome of newly infected EC cells and we could detect no progeny virus after infection of these cells, we wished to examine whether these proviral sequences were transcriptionally active although blocked with regard to virus production. Figure 5A,B shows that a small amount of full length retroviral RNA and envelope (3.5 kb) mRNA was detected in the poly(A)⁺ cytoplasmic RNA during the first 1 or 2 days post-infection and then it disappeared. In comparison, hybridization to the poly(A)⁺ RNA from infected differentiated cells was completely overexposed, as in Fig. 2B, lane g, and is not shown. Transcription rates were also measured by pulse-labelling isolated nuclei of infected EC cells at daily intervals post-infection to determine whether a burst of transcription occurred during the first few days after infection, producing primary transcripts that remain unprocessed, thus failing to appear as cytoplasmic poly(A)⁺ RNA. The results in Fig. 5C show that little or no proviral transcription was detectable in EC cells during the first 2 days post-infection, whereas in the infected differentiated virus-producing cells, proviral transcription occurred at a high rate. It is estimated that the level of retroviral poly(A)⁺ RNA present in EC cells 24 and 48 h post-infection represents 1–2 copies per cell. This RNA could also be from a rare differentiated cell that produces a large number of copies or it could represent transcription from endogenous proviruses. However, if it is endogenous retroviral RNA, then one must accept that its expression was transiently induced by M-MuLV infection in the absence of expression of M-MuLV provirus. In any case, whether the expression measured was from M-MuLV or endogenous provirus, the amount was insignificant relative to levels produced after infection of differentiated permissive cells. Moreover, this low level of retroviral RNA decreased after

30–40 h of infection while newly integrated M-MuLV provirus persists unmethylated for at least 1–2 weeks.

Conclusion

These results, taken together, show that in EC cells, provirus is unmethylated at CCCGGG sequences for at least 8 days after infection and yet is also not transcribed to any significant extent. Table 2 shows that DNA isolated from EC cells 5 days post-infection produced infectious M-MuLV in one of three cultures after transfection to NIH 3T3 cells. This supports the contention that M-MuLV provirus that is not methylated in CCCGGG sites is potentially infectious, whereas the hypermethylated provirus from the maintained BMBcl₂ cell line is not. We propose, therefore, that the primary cause of negative proviral expression in newly infected EC cells is due to a mechanism other than methylation of proviral CpG sequences. This does not agree with the conclusions of Stewart *et al.*⁹ that suppression of M-MuLV expression in F9 cells is correlated with methylation of *Sma*I sites on integration. It is possible, however, that immediate hypermethylation of a specific sequence not covered by *Sma*I recognition sites does occur and leads to negative provirus expression in EC cells but does not inhibit gene expression in NIH 3T3 cells. With regard to this alternative explanation, note that two of the *Sma*I recognition sites examined here, generating the 5' end of the 5.7-kb and the 3' end of the 2.1-kb fragments, reside in the repeat region of the M-MuLV LTR at genomic positions 32 and 8,294 (ref. 12). Retroviral LTRs have known transcriptional promoter activity^{24,25}. Thus, if methylation at multiple localized sites is required to limit gene expression, one might expect the methylation to occur within the promoter region, and to protect the CCCGGG sequence from *Sma*I cleavage.

Introduction of exogenous promoters into EC cells may trigger an immediate host-derived inhibition of transcription initiated at these new promoters, the latter stage of which involves methylation. Introduction of these promoters into differentiated cells, on the other hand, does not elicit the same expression-negative response by the host cell. Alternatively, tissue-specific transcription factors absent from EC cells may be required for proviral gene expression. An example of such a host factor has recently been described for poliovirus²⁶. It is possible that genes that are not transcribed in EC cells are recognized as such and hypermethylated to 'fix' the quiescent condition and that before they can be expressed once again they must be demethylated. During the preparation of this paper, several examples demonstrating an apparent lack of correlation between transcriptional activity and the pattern of methylation in the sequence of specific transcription units have been published^{27–29}. These observations further support our notion that hypermethylation alone is not the signal for gene regulation but may, in fact, be a consequence of gene inactivity during early embryogenesis.

This work was aided by a Basil O'Connor starter research grant (5-358) from March of Dimes Birth Defects Foundation. This is publication no. 2620 from the Research Institute of Scripps Clinic.

Received 28 June; accepted 6 October 1982.

- Jaenisch, R., Fan, H. & Croker, B. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4008–4012 (1975).
- Peries, J., Alves-Cardoso, E., Canivet, M., Devons-Guillemain, M. C. & Lasneret, J. J. *natn. Cancer Inst.* **59**, 463–465 (1977).
- Martin, G. R. *Science* **209**, 768–776 (1980).
- Teich, N., Weiss, R. A., Martin, G. R. & Lowy, D. R. *Cell* **12**, 973–982 (1977).
- Gautsch, J. W. *Nature* **285**, 110–112 (1980).
- Gautsch, J. W. *Somatic Cell Genet.* **8**, 143–149 (1982).
- Speers, W. C., Gautsch, J. W. & Dixon, F. J. *Virology* **105**, 241–244 (1980).
- D'Auriol, L. *et al. J. gen. Virol.* **55**, 117–122 (1981).
- Stewart, C. L., Stuhlman, H., Jähner, D. & Jaenisch, R. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4098–4102 (1982).
- Chattopadhyay, S. K., Lander, M., Rands, E., Lowy, D. R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5774–5778 (1980).
- Chan, H. W. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 5779–5783 (1980).
- Shinnick, T. M., Lerner, R. A. & Sutcliffe, J. G. *Nature* **293**, 543–548 (1981).
- Van Beveren, C., Goddard, J. G., Berns, A. & Verma, I. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3307–3311 (1980).
- Yoshimura, F. K. & Weinberg, R. A. *Cell* **16**, 323–332 (1979).
- Thomas, P. S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201–5205 (1980).
- Harpold, M. M., Evans, R. M., Salditt-Georgieff, M. & Darnell, J. E. *Cell* **17**, 1025–1035 (1979).
- Lerner, R. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **69**, 2965–2969 (1972).

- Levy, D., Lerner, R. A. & Wilson, M. C. *Proc. natn. Acad. Sci. U.S.A.* **79**, 5823–5827 (1982).
- Fan, H. & Verma, I. M. *J. Virol.* **26**, 468–478 (1978).
- Graham, F. L. & van der Eb, A. J. *Virology* **52**, 456–467 (1973).
- Stuhlmann, H., Jähner, D. & Jaenisch, R. *Cell* **26**, 221–232 (1981).
- Harbers, K. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 7609–7613 (1981).
- Youssefian, H. & Mulder, C. J. *molec. Biol.* **150**, 133–136 (1981).
- Neel, B., Hayward, W. S., Robinson, H. L., Fang, J. & Astrin, S. M. *Cell* **23**, 323–334 (1981).
- Fuhrman, S. A., Van Beveren, C. & Verma, I. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5411–5415 (1981).
- Dasgupta, A., Hollingshead, P. & Baltimore, D. *J. Virol.* **42**, 1114–1117 (1982).
- Weintraub, H., Beug, H., Groudine, M. & Graf, T. *Cell* **28**, 931–940 (1982).
- McKeon, C., Ohkubo, H., Pastan, I. & deCrombrughe, B. *Cell* **29**, 203–210 (1982).
- Macleod, D. & Bird, A. *Cell* **29**, 211–218 (1982).
- Nevins, J. R. & Wilson, M. C. *Nature* **290**, 113–118 (1981).
- Gautsch, J. W., Bedigian, H. G. & Meier, H. *Virology* **83**, 462–466 (1977).
- Jakob, H., Boon, T., Gaillard, J., Nicolas, J. F. & Jacob, F. *Ann. Microbiol. Inst. Pasteur* **124B**, 269–282 (1973).
- Stevens, L. C. *Adv. Morph.* **6**, 1–31 (1967).
- Speers, W. C., Birdwell, C. R. & Dixon, F. J. *Am. J. Path.* **97**, 563–578 (1979).
- Rigby, P. W. J., Diekmann, M., Rhodes, C. & Berg, P. J. *molec. Biol.* **113**, 237–251 (1977).
- McKnight, G. S. & Palmiter, R. D. *J. biol. Chem.* **254**, 9050–9058 (1979).

Unusual splice sites revealed by mutagenic inactivation of an authentic splice site of the rabbit β -globin gene

Berend Wieringa, Francois Meyer*, Jakob Reiser & Charles Weissmann

Institut für Molekularbiologie I, Universität Zürich, 8093 Zürich, Switzerland

Only one of six point mutations of the sequence around one end of the larger of the introns of the rabbit β -globin gene seriously affects the normal removal of the intron and splicing of the gene. That mutation converts a GT sequence, invariably found at the 5' end of introns, into an AT, which is no longer recognized as a signal for intron removal. Instead, three normally unused (cryptic) sites are used, leading to aberrant gene transcripts. One of the cryptic sites is an exception to the invariable GT sequence.

THE rule originally proposed by Breathnach *et al.*¹, which states that intron sequences invariably start at the 5' end with GT and end with AG (as underlined in the sequence below) has held true in all cases examined so far². More extensive consensus sequences have been derived²⁻⁶ for the region that extends on either side of the splice site at the junction of intron and exon:



where $n \geq 11$. In many 5' splice regions only six, and at least once only five, positions conform with the consensus sequence; the consensus run of 11 pyrimidines in the 3' splice region may contain as many as six purines, but never the doublet AG².

It has already been shown that deletion or inactivation of a splice site frequently does not abolish splicing, but leads to aberrant splicing in which sequences resembling a consensus structure (cryptic splice sites⁷) are used instead of their normal counterparts⁷⁻⁹.

Here we describe the preparation of a set of β -globin genes with point mutations in the consensus sequence of the 5' splice region of the large intron. Only one of six mutations abolished normal splicing, and led to the formation of three aberrant splice products. The sequences around the three cryptic 5' splice sites were determined; in one case the Breathnach-Chambon rule was violated.

Construction of rabbit β -globin genes with point mutations

Restructuring and mutagenesis were carried out on a 2.1-kilobase (kb) partial *Bgl*II fragment containing the rabbit β -globin ($R\beta$ G) gene¹⁰ (Fig. 1a), inserted in the *Bam*HI site of pBR322 (p $R\beta$ GA425B)¹¹, or pBR327¹².

Site-directed mutations were generated by the nucleotide-analogue substitution method¹³⁻¹⁵, with [N^4 -hydroxy]dCTP as the analogue (Fig. 2). In this technique nucleotide analogues are introduced into the sequence to be mutagenized by elongation of a primer-template complex using an appropriate combination of nucleotide analogue and standard nucleotides. The strand containing the nucleotide analogue at specific sites is then used as a template in a further *in vitro* reaction, and misincorporation opposite one of the analogues generates the required mutations.

To provide the single-stranded template plus strand $R\beta$ G DNA was cloned into the bacteriophage fd derivative 177¹⁶ in the appropriate orientation. The primer used was a single-

stranded *Bgl*II-*Bam*HI minus-strand fragment of a deletion mutant, $R\beta$ GA21, whose 3' end hybridized to position 502, eight residues downstream from the 5' splice site, and which had an unpaired C residue at the deletion end point (Fig. 2b).

The template-primer complex was incubated with DNA polymerase I to digest off the unpaired C of the primer, and then elongated to a limited extent using [N^4 -hydroxy]dCTP and either dATP plus dTTP to incorporate the analogue instead of dCMP (Fig. 2b, step 1, reaction A), or dATP plus dCTP to incorporate hydroxy-dCTP instead of dTTP (Fig. 2b, step 1, reaction B). The primer was then further extended with an excess of the four dNTPs, and separated from the template. The plus strand complementary to the [N^4 -hydroxy]dCTP-containing elongated minus strand was synthesized using an appropriate plus strand primer (Fig. 2, steps 4 and 5). The second strand was synthesized *in vitro* to generate the transition mutation before transfection, as the repair systems in *Escherichia coli* might remove the nucleotide analogue *in vivo*, thereby reducing the yield of mutants^{14,15}. The 641-base (bp) *Eco*RI-*Bam*HI fragment containing the mutagenized region was inserted into an $R\beta$ G-containing hybrid from which the homologous *Eco*RI-*Bam*HI fragment had been excised. The hybrid DNA was cloned, and 30 clones were sequenced; 13 had a single nucleotide change in the mutagenized region and no double mutants were observed. Six of the seven expected single point mutations were identified (Fig. 3b); only the G $\rightarrow$ A transition at position 499 was not obtained.

Expression of wild-type and modified rabbit β -globin genes in HeLa cells

Our expression vector (Fig. 1b) consisted of pBR327 (minus the pBR sequence deleterious for replication in eukaryotic cells¹⁷), a 3,039-bp segment of simian virus 40 (SV40) DNA with the enhancer sequence as well as the origin of replication and the complete early transcription unit (compare refs 18-20), and the mouse β -globin ($M\beta$ G) gene^{10,21} as an internal reference for the quantitation of transcription. The test gene, contained within a 2.7-kilobase (kb) *Cla*I-*Sal*I fragment, was inserted in the desired orientation into the expression plasmid (Fig. 1b). The transcription initiation sites of the reference and the test gene are about 2 kb apart, and the genes are divergently transcribed to minimize reciprocal interference. Details regarding the construction of this vector are available on request (B.W. and C.W.).

DNA was introduced into HeLa cells by the calcium phosphate method^{22,23} and total cellular RNA was purified after 48 h. S₁ mapping experiments showed that wild-type rabbit and mouse β -globin genes were transcribed from their own

*Present address: Ciba-Geigy AG, CH-4000 Basel, Switzerland.

promoters and processed normally (ref. 7 and additional data, not shown).

Qualitative and quantitative assay of β -globin transcripts in transfected cells

S_1 mapping, carried out with uniformly labelled rabbit and mouse minus-strand probes of the same specific activity⁷ revealed no cross-hybridization in the stringent annealing conditions used. Simultaneous hybridization of the probes with authentic mouse and rabbit β -globin mRNA (either separate or combined), followed by S_1 treatment, yielded the signals expected from the probes used⁷ (Fig. 1a): fragments of 144 and 145 nucleotides (not resolved) from the first exons of the mouse and rabbit β -globin transcripts respectively; a 222-nucleotide fragment from the (complete) second exon of the rabbit, and a 179-nucleotide fragment from (part of) the second exon of the mouse transcripts; and a 133-nucleotide fragment from (part of) the third exon of the rabbit transcript (data not shown).

In each experiment the ratio of radioactivities of the R β G-specific 222- or 133-nucleotide bands and the M β G-specific 179-nucleotide bands was determined (Fig. 3). The wild-type rabbit gene, irrespective of its orientation, gave rise to an approximately 2.5-fold higher transcript level than the mouse gene. If the SV40 fragment in the standard expression vector was inverted (clockwise late SV40 transcription, Fig. 1b, top) then the ratio of rabbit to mouse transcripts was significantly increased (data not shown).

Whatever the reason for the unequal ratio of transcripts, the expression vector used provides a valid reference system for the measurement of transcripts from modified genes.

The effect on splicing of point mutations in the 5' splice site region

Purine transitions in the first, second or third position (494, 493 or 492) upstream, or in the third or fourth position (497 or 498) downstream, of the splice site affected neither the quantity nor the structure of β -globin mRNA formed (Fig. 3).

The β -globin gene with a G to A transition at position 495, immediately downstream of the splice site, gave three aberrantly and no correctly spliced β -globin specific RNAs, at overall levels approximately one-half that of the wild-type control. S_1 mapping with the uniformly labelled probe (data not presented) gave the correct signals for exons 1 and 3, but the signal for exon 2 was missing. Three new protected fragments appeared instead: one was slightly longer than the 222-nucleotide fragment expected from exon 2; the second was approximately 170 nucleotides long; and the third 88 nucleotides long.

The transcripts were further characterized by downstream S_1 mapping with an 893-bp *TaqI*-*BglII* fragment 3'-terminally labelled at the *TaqI* site (position 308). The autoradiogram (Fig. 4A, lane a) showed three signals (or signal multiplets) corresponding to (1) 190-195, (2) 130-135 and (3) 50-52 nucleotide fragments. These findings indicate the existence of

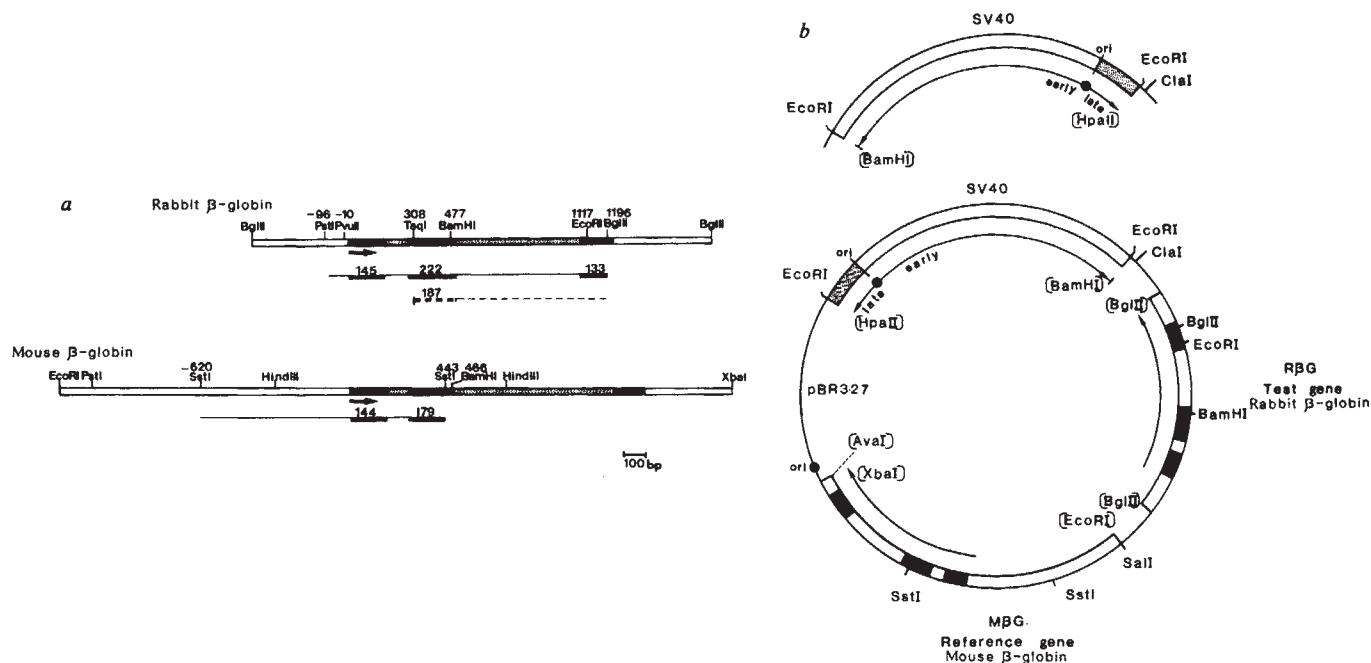


Fig. 1 Hybrid plasmid for the expression of modified rabbit β -globin genes. *a*, Structures of the rabbit β -globin construct used to test the effects of mutagenesis and the mouse β -globin construct that served as a control. Exons are shown as solid boxes, introns as stippled boxes and flanking sequences as open boxes. Only relevant restriction sites are shown (see ref. 10). Uniformly labelled probes (solid lines) and end-labelled probe (broken line), used for S_1 mapping as described in the text, are given below the genes. The heavy lines and the numbers next to them indicate the fragments protected by wild-type RNA. Arrows indicate direction of transcription. *b*, Structure of expression plasmids consisting of mouse β -globin chromosomal DNA (M β G, 3.1 kb), SV40 DNA in either clockwise (bottom), or anticlockwise (top) orientation (3.05 kb), wild-type or mutant rabbit β -globin chromosomal DNA (R β G, 2.1 kb) and a pBR327 fragment (*SalI*-*AvaI*, 2,509 bp). Sites obliterated by the cloning steps are in parentheses. Arrows indicate the directions of transcription.

Methods: The 3.1-kb *EcoRI*-*XbaI* fragment encompassing the wild-type mouse major β -globin gene²¹, obtained from plasmid M chr β G-36³⁹, was blunted with S_1 nuclease and purified. pBR327 was cut with *SalI* and *AvaI*, the ends filled in with T₄ DNA polymerase and the four dNTPs⁴⁰, and dephosphorylated. The 2,509-bp fragment carrying the *bla* gene was joined to the M β G DNA fragment. Clones containing the M β G gene in a clockwise orientation were identified. The *SalI* site at the junction between the 5' end of the M β G DNA and the pBR327 fragment was restored. An SV40-pBR327 recombinant plasmid (from Dr W. Schaffner) was cut with *HpaII* and *BamHI*⁴¹, the ends blunted with S_1 nuclease and the fragment containing the complete SV40 early region, origin and start of the late region was inserted into the filled-in *EcoRI* site (pBR position, 4,362/0) of the M β G-pBR327 recombinant. After transfection and cloning, hybrids with reconstituted *EcoRI* sites at both ends of the SV40 insert were identified. Constructs having SV40 DNA in the two possible orientations were isolated. M β G-SV40-pBR327 DNA was digested with *ClaI* and *SalI*; the large fragment served as acceptor for *ClaI*-*SalI*-excised R β G DNA. The β -globin DNA was either wild-type DNA or DNA modified by site-directed mutagenesis as described in Fig. 2 legend. Plasmid DNAs were prepared⁴² and characterized by restriction analysis and sequencing.

three cryptic splice sites, mapping approximately to positions 360, 440–445 and 497–499.

Primer extension analysis with a 5'-³²P-labelled minus-strand primer extending from position 1,120 to 1,068 gave three fragments of 285, 360–370 and 420–430 nucleotides, the latter (Fig. 4B, lane b) being slightly longer than the wild-type cDNA (Fig. 4B, lane a), and in the approximate ratio 5:2:3. There were no additional aberrant transcripts which might have been overlooked in the S₁ mapping experiment, in particular none shorter than 230 nucleotides, which would be expected if the third exon had been joined to a position upstream of the *TaqI* site in the second exon.

Localization and structure of the cryptic 5' splice sites

To determine the sequence and the precise position of the cryptic splice sites, we cloned and sequenced the relevant segments of double-stranded cDNAs of the three aberrant β -

globin RNAs (Fig. 4). Double-stranded cDNA prepared from poly(A)⁺ RNA of transformed cells²⁴ was digested with *TaqI* (position 308, R β G exon 2) and *BglII* (position 1,197, R β G exon 3), and fragments corresponding to the three size classes of aberrant globin-transcript copies, or to wild-type globin cDNA, were joined to the large *Clal*-*BamHI* fragment of pBR 327 and the hybrids cloned. One clone of each size class was sequenced (Fig. 4C), and the precise positions to which the third exon had been joined were determined.

As shown in Fig. 5a, the three types of abnormal transcripts resulted from splicing at cryptic splice sites 1, position 498/499 (...TGA/GTTTGG...); 2, position 442/443 (...TAA/GCTGAG...); and 3, position 359/360 (...AAG/GTGAAG...). Cryptic splice site 3 was the only one observed to be used in the case of R β G DNA in which a deletion from position 438 to 502 (R β G Δ 59) removed the authentic 5' splice site⁷. In a mutant with a deletion from position 482 to 502 (R β G Δ 21), cryptic site 2 was also used to some extent⁷.

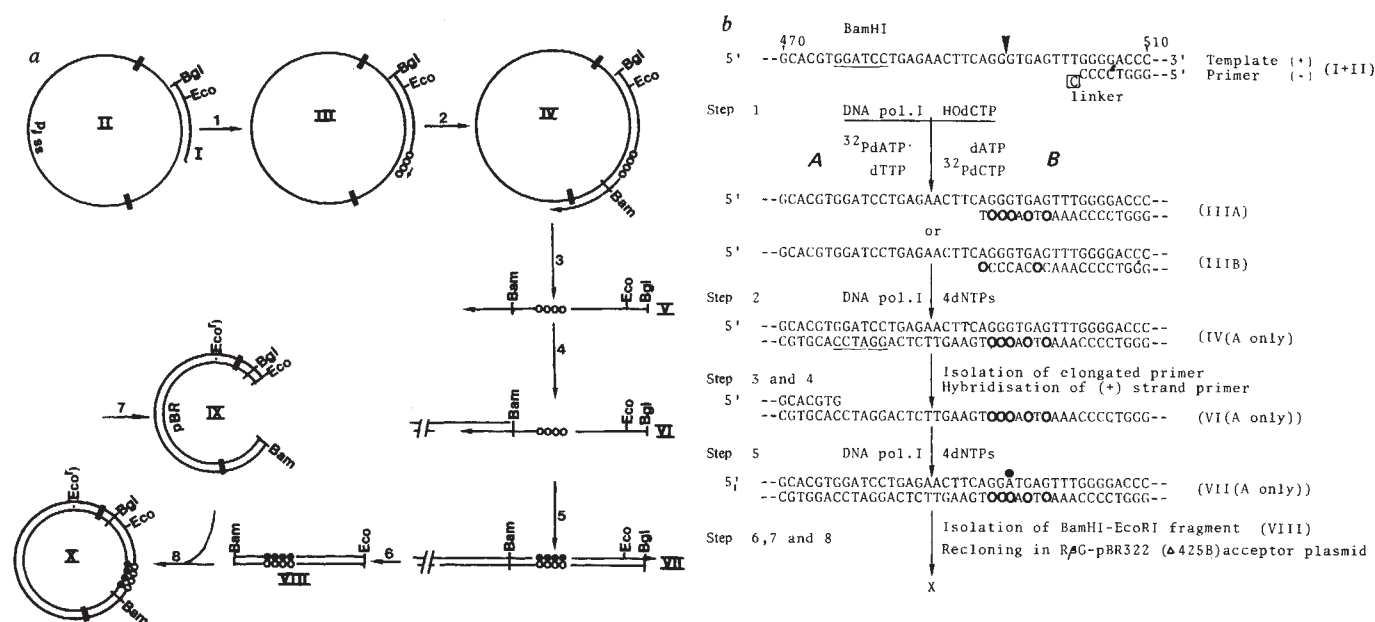


Fig. 2 Generation of point mutations in the 5' splice region of the large intron of the R β G DNA. *a* Shows, schematically, the eight steps involved in generation of the mutants. *b*, Details of changes in the sequence between positions 470–510 around the 5' splice site (arrow); open circles indicate sites of potential [hydroxy]dCTP incorporation for reaction A, the closed circle represents a mutation (G→A₄₉₅) generated during the synthesis of the second strand.

Methods: Plasmid R β G Δ 21 DNA contains the R β G sequence in which the DNA between positions 482 and 502 was replaced by a *BamHI* linker⁷. It was cleaved with *BamHI* and *BglII*, and the 700-bp fragment (I) purified on a 5–23% sucrose gradient; 2 μ g of primer DNA (I) were 5-terminally labelled with polynucleotide kinase and [γ -³²P]ATP, denatured at 100 °C in 2.5 mM EDTA, chilled and mixed with 107 μ g of single-stranded Z.fid177/Rchr-p Δ 425(+) phage DNA (II). This DNA contained the plus strand corresponding to the 2,071-bp partial *BglII* fragment of chromosomal R β G DNA¹¹ (see Fig. 1a), and was prepared by inserting pR β G Δ 425B DNA¹¹ into the *PstI* site of the replicative form of fd177¹⁶ (M. Kappeler, unpublished). Template II and primer I were hybridized in 1 ml 10 mM Tris-HCl (pH 8), 0.5 M NaCl for 1 h at 68 °C. After ethanol precipitation, the hybrid was separated from non-hybridized primer by sucrose gradient centrifugation. Step 1: the reaction mixture (100 μ l), with primer-template complex (I+II; containing 0.1 μ g of primer I), 50 mM Tris-HCl (pH 8.0), 5 mM MgCl₂, 10 mM β -mercaptoethanol, 50 μ g ml⁻¹ bovine serum albumin (BSA), 50 μ M [hydroxy]dCTP (given by Dr H. Weber) was adjusted to either 50 μ M dTTP, 20 μ M ³²P-dATP (*b*, reaction A) or to 50 μ M dATP, 20 μ M ³²P-dCTP (*b*, reaction B). Each mixture was incubated with DNA polymerase I (16 units; Biolabs) for 30 min at 20 °C. Step 2: 80 μ l of the reaction mixtures (III) were adjusted to 150 μ l 0.1 M K-phosphate (pH 6.9), 6 mM MgCl₂, 2 mM dithiothreitol (DTT), and 200 μ M each of dATP, dTTP, dCTP and dGTP, and 8 units of DNA polymerase I were added. After 90 min at 15 °C, EDTA was added to 10 mM and the samples were extracted with phenol and chloroform. Step 3: the samples (IV) were adjusted to 0.2 M NaOH and centrifuged through an alkaline sucrose gradient (4.2–19.2% sucrose in 0.8 M NaCl, 0.2 M NaOH, 10 mM EDTA) to separate extended primers (V) from fd template DNA and non-extended primer DNA. Fractions containing ³²P-labelled elongated primer (V) were pooled and the DNA was precipitated. Step 4: a plus strand primer for the second strand synthesis, prepared from *BamHI*-cleaved, 5' λ -exonuclease treated R β G DNA, was heat-denatured in 45 μ l 2.5 mM EDTA and hybridized as above. Step 5: annealed samples (VI) were incubated with DNA polymerase I (8 units) in 20 μ l 0.1 M K-phosphate (pH 6.9), 10 mM MgCl₂, 10 mM DTT, 0.4 mM each of dATP, dTTP and dGTP and 0.15 mM ³²P-dCTP at 15 °C for 6 h. The DNA was chromatographed on Sephadex G-50 and ethanol-precipitated. Step 6: the elongated, mutagenized DNA (VII) was digested with *BamHI* and *EcoRI*, extracted with phenol and ethanol-precipitated. Step 7: plasmid R β G Δ 425B (*EcoRI*: the *EcoRI* site in the pBR322 moiety was destroyed) was digested with *BamHI* and *EcoRI* and the large fragment (IX) isolated from a sucrose gradient. Step 8: approximately 10 ng of the 641-bp fragment (VIII) was joined to 60 ng of acceptor plasmid (IX) with T₄-DNA ligase. Cloned DNAs (X) were cleaved with *BamHI* and sequenced⁴³.

We were surprised that in the case of the 5' splice site in position 442/443, splicing occurred 5' to a GC doublet, rather than to the GT found in all other cases so far examined². It is unlikely that the DNA used for transfection contained 20–30% of a mutated species having GT at that position, as it had been recloned three times before sequencing and transfection. The cDNA sequence of the longest aberrant transcript allowed us to confirm the sequence around position 443 (arrow 2 in Fig. 4C). Moreover, the sequence TAAGCTGAGT, but not TAAGGTGAGT, contains an *AluI* site (underlined), and >95% of the template DNA was cleaved by *AluI* in this position (data not shown). Finally, a second cDNA clone corresponding to RNA spliced at the 442/443 position was found to be indistinguishable from the first by restriction mapping, showing that we had not fortuitously picked a non-representative cDNA clone.

Conclusion

Deletion or modification of a 5' splice region can lead to: (1) accumulation of unspliced precursor RNA, as reported for mutant adenovirus 5 EIA transcripts²⁵ and yeast actin precursor RNA²⁶; (2) retention of incompletely spliced precursors in the nucleus, as found for albumin mRNA of albuminaemic rats²⁷; or, (3) complete absence of transcripts^{28–30}. Alternatively, 5' splice sites normally used in other contexts^{8,31,32}, or cryptic splice sites^{7–9,33} can be utilized instead, leading to the appear-

ance of aberrantly processed RNA products, as in the results described above.

The compilation of splice regions⁴ clearly shows that a large variety of sequences are compatible with splicing. Our results show that the consensus sequence around one and the same splice site, the 5' splice region of the large rabbit β -globin intron, can be modified by point mutations (all base transitions) at various positions without detectably affecting the quality or efficiency of splicing; however, the G $\rightarrow$ A mutation converting the 'Chambon rule' GT to AT completely abolished splicing at this site. The same observation has been made for the natural GT $\rightarrow$ AT mutant of the human β -globin gene⁸ and for a GT to GG change generated *in vitro* at the 5' splice site of EIA RNA of adenovirus 5²⁸.

It has been suggested that a specific base-pairing interaction^{4,5,34} between small nuclear RNAs and the 5' and 3' terminal regions of an intron are part of the mechanism that mediates accurate splicing. The five base transitions that do not affect splicing do not disrupt the suggested base pairing. In that regard, our findings with the point mutations do not contradict the model.

Figure 5b shows a comparison of the wild-type and the cryptic 5' splice regions described in this paper with the corresponding 9-nucleotide consensus sequence². Whereas the sequences around the normal 5' splice site of the large β -globin intron, and the cryptic splice site 3 conform in seven and six positions, respectively, with the consensus sequence, the regions around

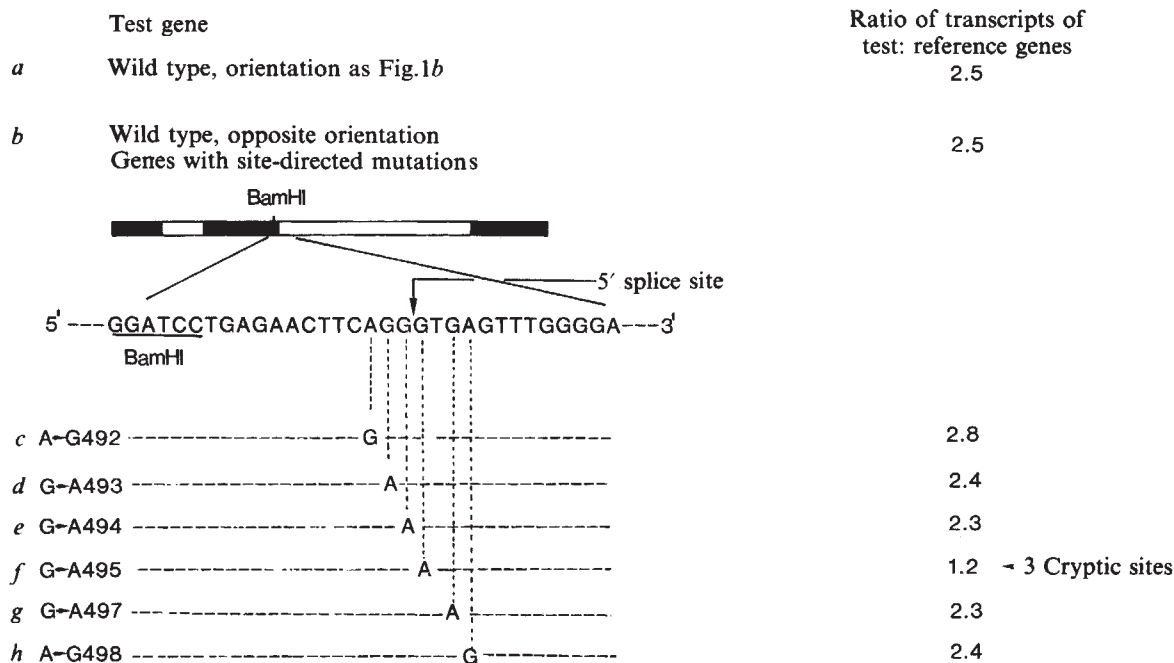


Fig. 3 Quantitation of transcripts of modified rabbit β -globin genes. The numbers on the right-hand side give the values for the level of transcripts of the test rabbit β -globin gene (shown on the left-hand side) relative to that of the test mouse β -globin gene.

Methods: Expression plasmid (12 μ g) bearing a wild-type or mutant rabbit β -globin gene (see Fig. 1) was mixed with 8 μ g sheared salmon sperm DNA and introduced into HeLa cells (9 cm dish, 50% confluent) by the Ca-phosphate method^{22,23}. The cells were shocked with dimethyl sulphoxide (DMSO)⁴⁴ 4 h after DNA application. After 40 h at 37 °C, total RNA was prepared by the LiCl-urea method⁴⁵. Purified probes (0.01 pmol of each) and RNA from transformed HeLa cells (10 μ g) were hybridized and used for S₁ mapping⁷. Transcripts were quantitated and analysed using uniformly labelled probes. To prepare R β G- and M β G-specific probes, the expression plasmid (Fig. 1b) was linearized with *SalI* and the DNA bearing the mouse and rabbit β -globin genes at its terminal regions was treated with exonuclease III to expose the plus strands of both genes. The DNA was repaired with [α -³²P]-dNTPs using DNA polymerase I and cut with *BglII*-*PstI* and *SstI* to yield probes specific for: (1) the first two exons and part of the third exon of rabbit β -globin RNA; and (2), for exon 1 and the greater part of exon 2 of mouse β -globin RNA, respectively (see Fig. 1a). S₁-mapping of β -globin mRNAs⁴⁶ with these probes yielded protected fragments of 145, 222 and 133 nucleotides for rabbit β -globin mRNA and of 144 and 179 nucleotides for mouse β -globin mRNA. There was no cross-hybridization between rabbit and mouse nucleic acids (see ref. 47). The ratio of test to reference transcripts (relative transcript level, or RTL) was calculated from the radioactivities of the 222-nucleotide fragment (R β G specific) and the 179-nucleotide fragment (M β G-specific) excised from the gel, or, where the second exon of R β G was shortened (sample f), from the radioactivity in the (R β G-specific) 133-nucleotide band relative to the 179-nucleotide mouse fragment, multiplied by 1.7 to correct for the difference in length of the 222- and 133-nucleotide fragments. Numbers given are the average of two independent experiments.

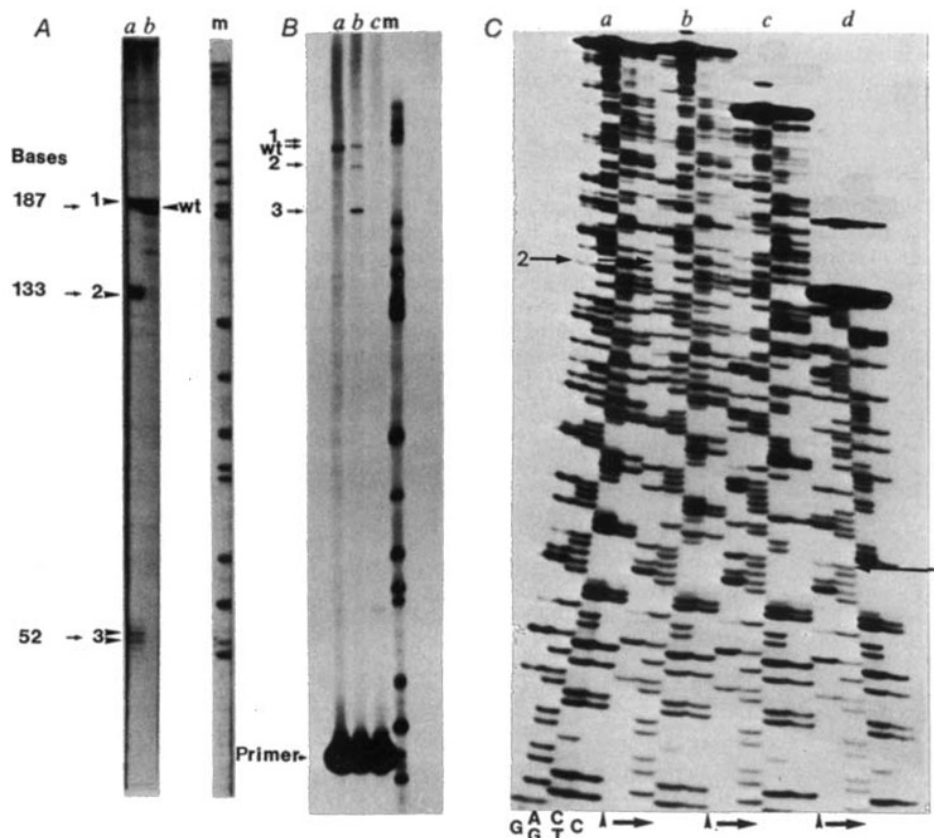
cryptic splice sites 1 and 2 conform at only three and two positions, respectively, values that might be expected from random sequences. Moreover, when one attempts to base pair the sequences downstream from the cryptic splice sites 1 and 2 with the 5' terminus of U1 RNA, as suggested by Lerner *et al.*⁵, no significant pairing is achieved (Fig. 5c).

In the case of cryptic splice site 2, where the Chambon 'GT rule' is formally broken, a better fit is obtained if the C residue

in the second position downstream from the splice site is looped out, which brings a U(T) residue next to the G, reconstituting, as it were, the 'Chambon rule' GT sequence, and allowing five base pairs to be formed.

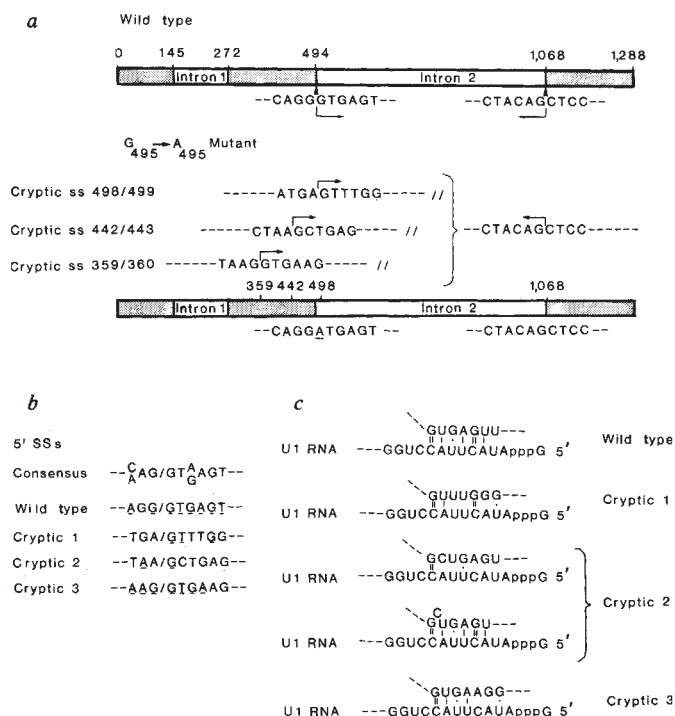
Inspection of the β -globin RNA sequence shows that there are many potential splice sites, as defined by the Steitz-Lerner model, especially if 'looping-out' is permitted. For example, at positions 394-404 in the second β -globin exon, the sequence

Fig. 4 Characterization of the incorrectly spliced transcripts of various modified R β G genes. **A**, S₁ mapping of the aberrant splice junctions. The RNA mapped in lanes *a* and *b* was prepared from cells transfected with plasmid vector carrying either a G₄₉₅→A₄₉₅ mutant R β G gene (lane *a*), or a wild type R β G gene (lane *b*). Lane *m* is a marker track with ³²P-labelled pBR322 BspI fragments. **B**, primer extension mapping of transcripts of the wild type (lane *a*) or G₄₉₅→A₄₉₅ mutant (lane *b*) R β G gene. Lane *c* is a control using RNA from mock-transfected cells, and lane *m* a marker track as in **A**. **C**, sequencing ladders of cloned globin cDNAs derived from correctly and aberrantly spliced β -globin poly(A)⁺ RNAs. The cloned cDNAs are derived from: wild-type R β G RNA (lane *a*); mutant R β G RNAs with cryptic splice sites at positions: 498/499 (lane *b*); 442/443 (lane *c*); 359/360 (lane *d*). In each case the sequence reads upstream from the 5'-labelled EcoRI end in the third exon, into the second exon. Arrow 1 indicates the 5' end of exon 3, and arrow 2 the position in lane *a* where the sequence of the cryptic splice region 2 can be read.



Methods: **A**, S₁ mapping. A 3'-labelled probe was prepared by filling in the *Taq*I site of the 893-bp *Taq*I-*Bgl*II fragment (see also Fig. 1a) with dGTP and [α -³²P]dCTP (specific activity 400 Ci mmol⁻¹) using the Klenow fragment of DNA polymerase I. Preparation of RNA from cells transformed with the rabbit β -globin-containing DNAs, and S₁ mapping, using 20 μ g of total RNA, were as described in Fig. 3 legend. **B**, primer extension mapping. A single-stranded *Alu*I-EcoRI primer fragment (minus strand, positions 1,068-1,120, 0.015 pmol) was 5'-³²P-labelled and hybridized in conditions described previously⁴⁸ to the poly(A)⁺ RNA prepared from either 150 μ g total RNA of HeLa cells transfected with wild-type R β G plasmid (lane *a*); from 60 μ g of total RNA of cells transfected with G₄₉₅→A₄₉₅ mutant DNA (lane *b*); or from 150 μ g total RNA of mock-transfected cells (lane *c*). After ethanol precipitation the hybrids were elongated in 20 μ l 50 mM Tris-HCl (pH 8), 6 mM dithiothreitol, 10 mM MgCl₂, 80 mM KCl and 500 μ M each of the four dNTPs with reverse transcriptase (15 units) for 1 h at 41 °C. The products were electrophoresed through a 6.5% acrylamide-urea sequencing gel and autoradiographed. The positions of full-length cDNAs of wild type (wt) and mutant transcripts (aberrant transcripts 1, 2 and 3) are indicated by arrows. **C**, cDNA sequencing. RNA (five times the amounts given above for the primer extension experiment) from HeLa cells transformed with wild-type R β G DNA or the G₄₉₅→A₄₉₅ mutant DNA was hybridized to oligo (dT)₁₅₋₁₈ and single-stranded cDNA synthesized with reverse transcriptase essentially as above. Double-stranded cDNA was prepared using DNA polymerase I and the reaction conditions described by Wickens *et al.*²⁴. RNA was degraded with RNase A (100 μ g ml⁻¹), the double-stranded cDNA purified on Sephadex G-50 and digested with *Taq*I and *Bgl*II. The fragments were separated on a 2% low melting agarose gel and fractions corresponding to cDNAs of 300-350 bp, 220-300 bp and 150-220 bp excised. pBR327 DNA was cut with *Cl*AI and *B*amHI, electrophoresed through a 0.8% low melting agarose gel and the large fragment ('acceptor DNA') excised. Each cDNA-containing fraction was joined to acceptor DNA with T₄ DNA ligase. The DNAs were transfected into *E. coli* HB101 and R β G-specific colonies were identified by colony hybridization⁴⁹ using the nick-translated *Pvu*II-EcoRI 1,180-bp fragment from R β G DNA (see Fig. 1a) as probe. Plasmid DNAs derived from one and the same cDNA fraction were identical as judged by *Alu*I and *B*spI restriction analysis. 5'-³²P-labelled EcoRI fragments (containing the R β G region of interest) of one representative clone of each class were isolated from an agarose gel and purified⁵⁰. The DNA was cleaved with *Taq*I (at position 23 of pBR327¹² and position 308 of R β G¹⁰) and the R β G-containing fragment sequenced according to Maxam and Gilbert (ref. 43; see also ref. 51).

Fig. 5 Cryptic splice sites revealed by mutational inactivation of the 5' splice site of the large β -globe intron. *a*, Location and sequence. Brackets with arrowheads indicate the regions (at the DNA level) that are spliced out from wild-type transcripts (above) or from transcripts of the gene with the GT $\rightarrow$ AT mutation (below). The numbers indicate map positions¹⁰. *b*, Comparison of wild-type and cryptic 5' splice regions with the 5' splice region consensus sequence²; matching nucleotides are underlined. *c*, Possible base-pairing of the 5' terminal region of snU1 RNA with the sequence around the wild-type or cryptic 5' splice sites. The 5' end of the snU1 RNA as well as seven nucleotides downstream of the correct or aberrant splice sites (starting at positions 499 (cryptic 1), 443 (cryptic 2), 360 (cryptic 3)) are paired in the same register as proposed by Lerner *et al.*⁵. In the case of cryptic 2, looping out of a C residue allows the same base-pairing pattern as proposed for the wild-type splice region.



..TCA/GTGAGGGT.. could stably base pair with U1 RNA, and thus represents as good a candidate for a cryptic splice site as any other, however no mRNAs were observed that had been spliced at that position. As the sequence is localized between two cryptic splice sites that are utilized, a scanning model⁶ cannot explain why it is not used while the other two are. It could be that other sequence elements are required; for example, the consensus sequence shows the prevalence of the sequence AAG preceding the 5' splice site. This requirement, which is not explained by the Lerner-Steitz model, is met neither by cryptic splice sites 1 and 2 nor by various natural splice sites².

Perhaps some sequences that potentially define a splice site are not used because they are masked by secondary structure, as is thought to be the case for some potential protein initiation sites in prokaryotes^{35,36}. In addition, the tertiary structure of the pre-mRNA may preclude the use of some combinations of splice sites, because they may be placed at spatially distant or unfavourable positions. Nevertheless, a specific tertiary structure is not required for splicing, as shown by the findings that hybrid RNAs, joined within an intron, are spliced normally³⁷, and that extensive deletions or modifications within introns or flanking exon regions (refs 7, 29, 38 and this paper) do not affect the accuracy of splicing.

We have found that if a 136-bp duplication of the β -globin 5' splice region of the large intron is inserted upstream of the original 5' splice site, then splicing occurs exclusively at the more upstream copy. A similar experiment, carried out with

the 3' splice region, has shown that the more downstream of the two splice sites is utilized (T. Kühne and B. Wieringa, unpublished results). This again shows that factors other than just the nucleotide sequences in the immediate neighbourhood of the splice sites determine the final splicing pattern, and that there is no unique relationship between the normally utilized 5' and 3' splice sites as regards their relative position on the pre-mRNA.

Thus, it seems likely that a potential splice site and its affinity for the splicing system are primarily defined by one of a set of nucleotide sequences, but the actual use of a pair of sites is modulated by the secondary and tertiary structure of the RNA, and possibly by proteins associated with it.

This work was supported by the Schweizerische Nationalfonds and the Kanton of Zürich; B.W. was the recipient of a long-term EMBO Fellowship. We thank Hans Weber and Peter Dierks for helpful advice.

Received 16 August; accepted 11 November 1982.

- Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4853-4857 (1978).
- Mount, S. M. *Nucleic Acids Res.* **10**, 459-471 (1982).
- Seif, I., Khoury, G. & Dhar, R. *Nucleic Acids Res.* **6**, 3387-3398 (1979).
- Rogers, J. & Wall, R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1877-1879 (1980).
- Lerner, M. R., Boyle, J. A., Mount, S. M., Wolin, S. L. & Steitz, J. A. *Nature* **283**, 220-224 (1980).
- Sharp, P. A. *Cell* **23**, 643-646 (1981).
- Wieringa, B., Meyer, F., Reiser, J. & Weissmann, C. *ICN-UCLA Symp. molec. cell. Biol.* (in the press).
- Treisman, R., Proudfoot, N. J., Shander, M. & Maniatis, T. *Cell* **29**, 903-911 (1982).
- Felber, B. K., Orkin, S. H. & Hamer, D. H. *Cell* **29**, 895-902 (1982).
- van Ooyen, A., van den Berg, J., Mantei, N. & Weissmann, C. *Science* **206**, 337-344 (1979).
- Dierks, P., van Ooyen, A., Mantei, N. & Weissmann, C. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1411-1415 (1981).
- Soberon, X., Covarrubias, L. & Bolivar, F. *Gene* **9**, 287-305 (1980).
- Flavell, R. A., Sabo, D. L. O., Bandle, E. F. & Weissmann, C. *Proc. natn. Acad. Sci. U.S.A.* **72**, 367-371 (1975).
- Müller, W., Weber, H., Meyer, F. & Weissmann, C. *J. molec. Biol.* **124**, 343-358 (1978).
- Weber, H. *et al. ICN-UCLA Symp. molec. cell. Biol.* **23**, 367-385 (1981).
- Herrmann, R., Neugebauer, K., Pirkl, E., Zentgraf, H. & Schaller, H. *Molec. gen. Genet.* **177**, 231-242 (1980).
- Lusky, M. & Botchan, M. *Nature* **293**, 79-81 (1981).
- Banerji, J., Rusconi, S. & Schaffner, W. *Cell* **27**, 299-308 (1981).
- Gruss, P. & Khoury, G. *Proc. natn. Acad. Sci. U.S.A.* **78**, 133-137 (1981).
- Fukushima, Y. *et al. Cell* **28**, 585-593 (1982).
- Tilghman, S. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **74**, 4406-4410 (1977).
- Graham, F. L., & van der Eb, A. J. *Virology* **52**, 456-467 (1973).
- Wigler, M., Pellicer, A., Silverstein, S. & Axel, R. *Cell* **14**, 725-731 (1978).

- Wickens, M. P., Buell, G. N. & Schimke, R. T. *J. biol. Chem.* **253**, 2483-2495 (1978).
- Carlock, L. & Jones, N. C. *Nature* **294**, 572-574 (1981).
- Gallwitz, D. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3493-3497 (1982).
- Esumi, H. *et al. Cold Spring Harb. Meeting on RNA Processing, Abstr.* 132 (1982).
- Montell, C., Fisher, E. F., Caruthers, M. H. & Berk, A. J. *Nature* **295**, 380-384 (1982).
- Khouri, G., Gruss, P., Dhar, R. & Lai, C.-J. *Cell* **18**, 85-92 (1979).
- Lai, C.-J. & Khoury, G. *Proc. natn. Acad. Sci. U.S.A.* **76**, 71-75 (1979).
- Choi, E., Kuehl, M. & Wall, R. *Nature* **286**, 776-779 (1980).
- Seidman, J. G. & Leder, P. *Nature* **286**, 779-783 (1980).
- Orkin, S. H. *et al. Nature* **296**, 627-631 (1982).
- Ohshima, Y., Itoh, M., Okada, N. & Miyata, T. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4471-4474 (1981).
- Lodish, H. F. *J. molec. Biol.* **50**, 689-702 (1970).
- Gheysen, D., Iserentant, D., Derom, C. & Fiers, W. *Gene* **17**, 55-63 (1982).
- Chu, G. & Sharp, P. A. *Nature* **289**, 378-382 (1981).
- Gruss, P. & Khoury, G. *Nature* **286**, 634-637 (1980).
- van den Berg, J. *et al., Nature* **275**, 37-44 (1978).
- Challberg, M. D. & Englund, P. T. *Meth. Enzym.* **65**, 39-43 (1980).
- Toozee, J. in *DNA Tumor Viruses* (ed. Toozee, J.) (Cold Spring Harbor Laboratory, New York, 1980).
- Wieringa, B., Ab, G. & Gruber, M. *Nucleic Acids Res.* **9**, 489-501 (1981).
- Maxam, A. & Gilbert, W. *Meth. Enzym.* **65**, 499-560 (1980).
- Lewis, W. H., Srinivasan, P. R., Stokoe, N. & Siminovich, L. *Somatic Cell Genet.* **6**, 333-348 (1980).
- Auffrey, C. & Rougeon, F. *Eur. J. Biochem.* **107**, 303-314 (1980).
- Berk, A. J. & Sharp, P. A. *Cell* **12**, 721-732 (1977).
- Weaver, R. F. & Weissmann, C. *Nucleic Acids Res.* **7**, 1175-1193 (1979).
- Nagata, S., Mantei, N. & Weissmann, C. *Nature* **287**, 401-408 (1980).
- Hanahan, D. & Meselson, M. *Gene* **10**, 63-67 (1980).
- Wieslander, L. *Analyt. Biochem.* **98**, 305-309 (1979).
- Smith, D. R. & Calvo, J. M. *Nucleic Acids Res.* **8**, 2255-2274 (1980).

Detection of 10^{14} -eV iron nuclei in cosmic rays

R. K. Sood

Physics (RAAF) Department, University of Melbourne, Parkville, Victoria 3052, Australia

Recent measurements of the charge and energy composition of cosmic rays have shown the energy spectra of secondary cosmic ray nuclei (Li, Be, B) to be steeper than those of the parent nuclei (C, O) indicating a decreasing path length for the particles with increasing energy. Differences also exist in the energy spectra of primary nuclei such as C, O and Fe. Of particular interest is the spectrum of Fe above 100 GeV. Measurements exist up to $\sim 10^{13}$ eV (refs 1, 2), and these indicate that the spectrum of Fe is less steep than that of lighter nuclei including protons. If this trend were to continue to $>10^{15}$ eV, Fe would then become the dominant constituent of cosmic rays. This conclusion is supported by results from air shower measurements at primary energies $>10^{13}$ eV (refs 3–5). However, air shower measurements suffer from the inherent uncertainty of having to extrapolate data from lower energy accelerator measurements for air shower modelling⁶. Thus these results need confirmation by direct measurements of Fe nuclei. We describe here results from such an experiment.

Conventional scintillator–Cerenkov type detectors flown on balloons have a typical geometric factor of $\sim 1 \text{ m}^2 \text{ sr}$. They therefore require unrealistically large exposure times to obtain meaningful statistics at 10^{14} eV where the total particle flux is $\sim 1 \text{ m}^2 \text{ h}^{-1}$. The experiment described here uses a different technique to overcome these problems. It relies on the detection of the light generated by relativistic nuclei in the upper atmosphere before disintegration, a technique first suggested by Sitte⁷ and discussed in detail by Gough⁸. A 10^{14} -eV Fe nucleus travelling at 45° to the vertical will have generated $\sim 1.2 \times 10^6$ photons from the top of the atmosphere to a balloon float altitude of 34 km. This light will be spread out over a radius, which at a constant altitude is a function of the particle energy and point of interaction in the atmosphere. For a 10^{14} -eV Fe nucleus which does not suffer an interaction the light pool radius is 13.6 m, yielding a photon density of $\sim 2,000$ photons m^{-2} . Such a flux can easily be detected above the intense night-sky background. The detector used in this experiment is shown in Fig. 1 and is described in ref. 9.

The main features of the detector are a pair of light-collecting dishes, each in the shape of a deep parabola, with a hemi-

spherical phototube at the focus. The inside surface of the dish which is made of fibre glass is polished and aluminized. The signals from the two phototubes I and II are used in coincidence to eliminate noise due to (1) the intense night-sky light and (2) cosmic-ray particle interactions which would masquerade as genuine events. The hemispherical phototubes were selected for a 2π sr response and the inherent thin glass envelope which minimizes spurious signal from the passage of cosmic rays. Five scintillators (Fig. 1) are also used in anticoincidence to eliminate cosmic-ray-generated background.

The threshold of the detector in photons m^{-2} is determined by the fluctuations in the noise level detected due to the night-sky background light. The minimum detectable light intensity is given by

$$\phi_c = s \left(\frac{\phi_B \Omega \tau}{\epsilon A} \right)^{1/2} \quad (1)$$

where s is the signal-to-noise ratio; ϕ_B , the photon flux from night-sky ($= 9.8 \times 10^7$ photons $\text{cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1}$); Ω , the field of view $= 0.054$ sr; τ , the resolving time $= 10^{-8}$ s; ϵ , the photocathode quantum efficiency $= 0.2$; A , collecting area of dish $= 1.1 \text{ m}^2$.

Substituting the relevant numbers in equation (1) yields $\phi_c = 350$ photons m^{-2} for $s = 4$, compared with $\sim 2,000$ photons m^{-2} from a Cerenkov light pulse from an iron nucleus. The pulse height from each tube is recorded for each coincidence detected. The energy threshold for the experiment is determined by the Cerenkov light yield which is a function of the varying refractive index of the air. The effective Cerenkov threshold at 34 km altitude is 2.4×10^{13} eV.

An important feature of the experiment is that the geometric factor is not limited by the physical size of the dish, but is given by the product of the mean solid angle and the area of the Cerenkov light pool. Values of $>30 \text{ m}^2 \text{ sr}$ are easily possible, depending on the float altitude.

The payload is tilted at 45° to the vertical to avoid the shadow of the balloon. The payload weighing ~ 400 lb was flown successfully from Alice Springs on 1 December 1981, using a $2.99 \times 10^6 \text{ ft}^3$ balloon. Launch was at 18.21 LT, the balloon attained a float altitude of 6.6 mbar and maintained this level to within ± 0.4 mbar. The flight was terminated after 10.5 h to avoid the balloon going over the horizon. The high voltage to the photomultipliers was switched on at 20.40 LT, ~ 1 h 40 min after sunset. This is the time required for the air glow to settle to a steady level after sunset: 6.5 h of useful data were recorded.

The phototube outputs and other key points in the payload electronics were monitored continuously to check on gain variations during the flight. A total of 16 coincidences was recorded

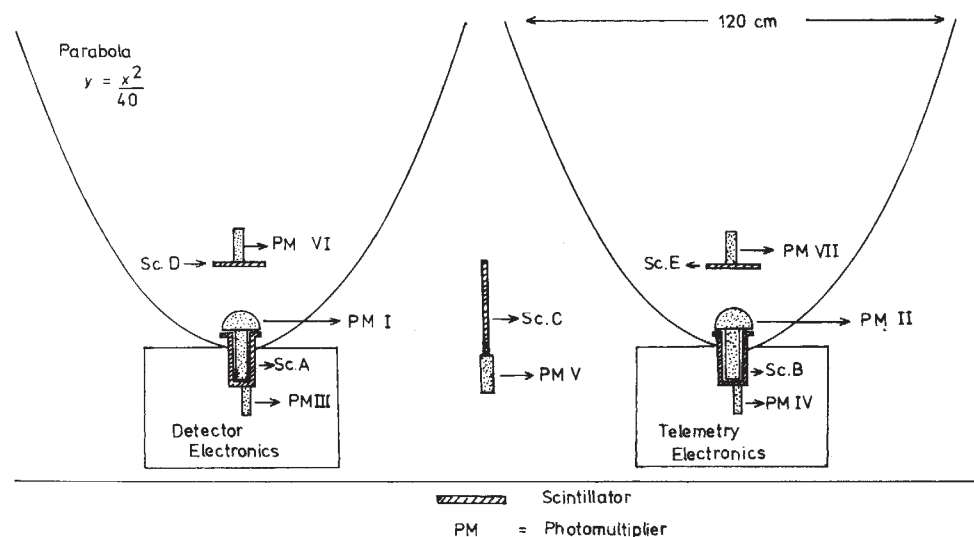


Fig. 1 Schematic of the detector.

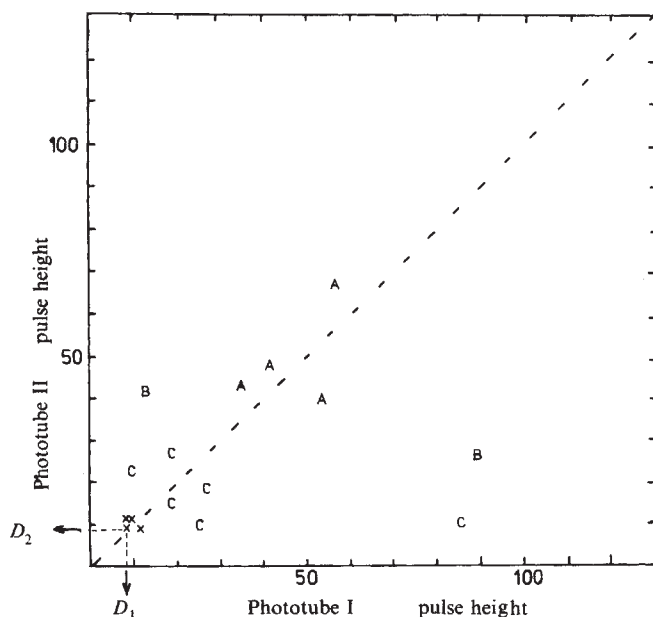


Fig. 2 Two-dimensional PHA plot: A, accepted iron nuclei; B, probable iron nuclei; C, possible lower energy iron, lower Z nuclei and background events; X, starlight background events.

during the flight. These are shown in a two-dimensional PHA plot in Fig. 2. This plot shows the output of phototube II against the output of phototube I. For an ideal detector, the points should be clumped around the 45° line, the scatter being due mainly to photoelectron statistical fluctuations. The charge resolution is then defined by this scatter. In the present case, the situation is made worse by the uneven Cerenkov light density from the axis of the particle trajectory⁸. Also, the photoelectron fluctuation which is a function of the photocathode to first dynode voltage, is large due to the low operating voltage necessary to limit the anode current in the face of the large night-sky background. Discriminator thresholds D_1 and D_2 in Fig. 2 correspond to 350 photons m^{-2} while pulse height 60 corresponds to an input flux of 2,000 photons m^{-2} . The calibration of this detector was carried out using: (1) a pulsed light source with a high ambient background light level to simulate the night sky; (2) a ^{241}Am source embedded in a scintillator; (3) observations of ground level air showers. Because of the inherent uncertainty in the above methods, the calibration figures could have an error of up to a factor of two.

The events recorded have been given different emphasis in Fig. 2, reflecting the certainty with which they can be attributed to incoming Fe nuclei. Considering the low number of total events recorded, it is meaningless to try to fit an expected computed distribution to that observed. Therefore, the following argument is presented to classify the events. A detailed analysis will be presented elsewhere. (1) The sub-Fe elements Sc, Ti, V, Cr and Mn have Cerenkov thresholds $>2 \times 10^{13}$ eV at the relevant altitude. The flux of these particles is expected to be $\sim 1\%$ of the Fe flux in view of the HEAO 3 results¹⁰ which exhibit a decreasing relative abundance to Fe with increasing energy at $0.5\text{--}30$ GeV nucleon^{-1} . Nuclei of $Z > 26$ indicate a similar scarcity¹¹. (2) The probability of recording background events, due to many particles simultaneously passing through the hemispherical tube structure or primary protons interacting in the atmosphere to produce ~ 600 secondaries can be shown to be negligible. The 'A' events in Fig. 2 are therefore designated as genuine Fe events. 'B' events could be due to Fe where the particle axis has passed close to one dish. The radial dependence of the photon density at 34 km shows a similar structure to that calculated by Gough for 40 km (ref. 8). (3) The low amplitude registered by the 'C' events could in principle be due to Fe nuclei which interact high in the atmosphere thus producing an annular ring of light of low photon

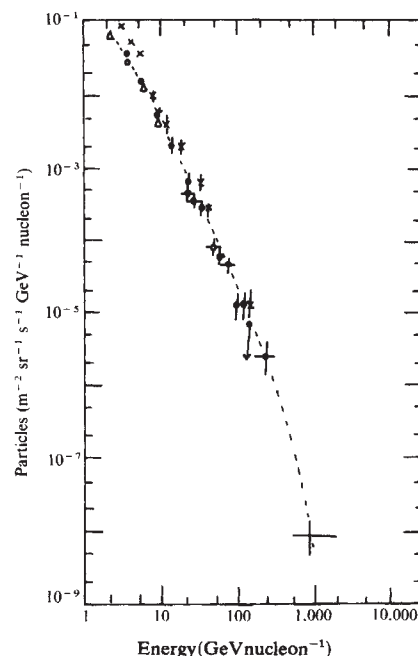


Fig. 3 Iron differential spectrum. Data sources: $\bullet$, ref. 12; $\times$, ref. 13; $\circ$, ref. 14; Δ , ref. 15; $+$, present work.

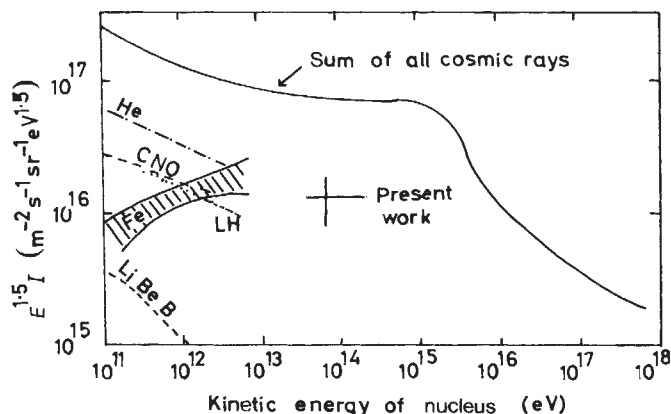


Fig. 4 Cosmic ray integral spectrum.

density at 34 km. Calculations show that $<5\%$ of the particles will interact to produce the required distribution. The probability of the payload being located at the edge of the light pool to produce C events is $<4\%$. We therefore attribute these events to sub-Fe nuclei produced in atmospheric interactions of Fe. The details of the types of the particles produced, and interaction cross-sections are unimportant, as in converting the A and B events to an unattenuated flux at the top of the atmosphere, a correction is automatically applied for the C events. (4) The 'X' events, very close to the discrimination thresholds are attributed to chance coincidences due to the night sky light, as their rate could be greatly increased by lowering the threshold levels. These events were common in ground level tests.

Based on measurements at lower energies of other workers Si should be abundant enough to be detectable by this technique, but the Cerenkov photon yield of ~ 250 photons m^{-2} is too close to the threshold to yield meaningful results.

To derive a flux at the top of the atmosphere for A + B events, corrected for atmospheric interactions, the geometric factor of the detector, which is a function of energy has been folded in with an assumed $E^{-2.3}$ Fe spectrum. The differential flux thus calculated is plotted in Fig. 3. The horizontal error bar corresponds to the energy limits where the detector response is down to 50%. Also shown are the results of other workers at lower

energies¹²⁻¹³. A steepening of the spectrum is apparent from the data points.

To accentuate the change in the slope the results are plotted on a different scale in Fig. 4 where the integral flux has been multiplied by a factor $E^{1.5}$. The data from previous work are taken from ref. 16. Clearly our data point falls below the extrapolation from lower energies.

Careful payload design and flight procedures should have eliminated extraneous sources of background in this experiment. Data analysis shows an expected randomness in arrival time and arrival direction. Lack of correlation with bright stars in the field of view has also been confirmed. We therefore have to conclude, on the basis of this limited amount of data that the primary Fe-spectrum steepens between $\sim 10^{13}$ and 10^{14} eV. The implications of this result, which we plan to confirm in 1983, are significant. Air shower results have indicated that Fe

Received 20 August; accepted 10 November 1982.

1. Simon, M. *et al. Astrophys. J.* **239**, 712–724 (1980).
2. Balasubrahmanyam, V. K. *Proc. 16th int. Conf. Cosmic Rays*, Kyoto **14**, 121–134 (1979).
3. Goodman, J. A. *et al. Phys. Rev. Lett.* **42**, 854–857 (1979).
4. Cowsik, R. *et al. Proc. 17th int. Conf. Cosmic Rays*, Paris **2**, 120–123 (1981).
5. Thornton, M. & Clay, R. J. *Phys. G5*, L137–L140 (1979).
6. Gassier, T. K. & Yodh, G. B. A. *Rev. nucl. part. Sci.* **30**, 475–542 (1980).
7. Sitte, K. *Proc. 9th int. Conf. Cosmic Rays*, 2, 887–889 (1965).
8. Gough, M. P. *J. Phys. G2*, 965–969 (1976).

Mixing motions in the Sun and solar-type stars

J. Nittmann

Department of Applied Mathematical Studies, University of Leeds, Leeds LS2 9JT, UK

W. P. Wood

Department of Mathematics, University of Newcastle, Newcastle, New South Wales, Australia 2308

We consider here some of the consequences of adiabatic dynamically driven motions within the radiative inner region of the Sun and solar-type stars for which it is assumed that an oblique magnetic field is trapped below the convection zone. The internal motions are due to both centrifugal and magnetic forces and we discuss their behaviour over the lifetime of a solar-type star. We assume, following Mestel¹, that stars of mass $\sim 1 M_{\odot}$ have a magnetic field which is contained by the outer convection zones of solar-type stars. We propose that if such a field exists in the interior of the Sun and other solar-type stars then coupled with the possibility of a rapidly rotating core as suggested by Claverie *et al.*², we have a possible mechanism for diffusive-type motions which could influence solar models³. Furthermore, we show that if the nutation frequency of an oblique magnetic rotator corresponds to the solar cycle then the magnetic field will need to be of the order of 10^7 G. All of the relevant theory and equations may be found in refs 4 and 5 and here we merely outline some of the pertinent assumptions and results contained therein.

In performing the calculations for a solar-type star we have assumed that the prescribed magnetic field is purely toroidal, symmetric about an axis which is inclined at an angle χ to the rotation axis of the star. This basic field is chosen for mathematical simplicity, however, such fields are subject to instabilities near the axis of symmetry⁶. In our choice of such a simple field we assume implicitly that a magnetic field exists whose geometry has the given complexity to make it stable. A further assumption that we make is that a more realistic magnetic field configuration will not greatly alter our order of magnitude estimates for the internal motions.

The magnetic and centrifugal forces are assumed small enough for a first-order double perturbation analysis to be valid. These forces, nevertheless, give the star three unequal axes of inertia so that the motion can be analysed into: (1) a uniform rotation; (2) an Eulerian nutation about the magnetic axis, with a frequency ω , fixed by the rotation of the star and the magnetic

should become the major cosmic-ray constituent at $\sim 10^{15}$ eV. This conflicting view may indicate that air shower models based on scaling from lower energies need some corrections. Our results also suggest that the acceleration and propagation models for heavy nuclei may need to be modified. For example, in the widely accepted nested leaky box model of Cowsik and Wilson¹⁷ all nuclear species are accelerated to the same spectral index in the source, and the spectra are modified at lower energies through interactions in interstellar space. The path lengths at higher energies are shorter with no consequent change in spectral index. The above results suggest that the cosmic ray confinement region may have smaller dimensions than previously accepted.

We thank J. Panettieri for technical help, Professor J. A. Thomas for support and helpful discussions and Dr G. K. Rochester for advice and encouragement.

9. Sood, R. & Panettieri, J. *Nucl. Instrum. Meth.* **185**, 427–431 (1981).
10. Koch, L. *et al. Proc. 17th int. Conf. Cosmic Rays*, Paris **2**, 18–21 (1981).
11. Klarmann, J. *et al. Proc. 17th int. Conf. Cosmic Rays*, Paris **2**, 13–16 (1981).
12. Simon, M. *et al. Proc. 16th int. Conf. Cosmic Rays*, Kyoto **1**, 352–357 (1979).
13. Orth, C. D., Buffington, A., Smoot, G. F. & Mast, T. S. *Astrophys. J.* **226**, 1147–1161 (1978).
14. Juliusson, E. *Astrophys. J.* **191**, 331–348 (1974).
15. Meyer, P. & Minnagawa, G. *Proc. 16th int. Conf. Cosmic Rays*, Kyoto **1**, 329 (1979).
16. Hillas, A. M. *Proc. 16th int. Conf. Cosmic Rays*, Kyoto **8**, 7–12 (1979).
17. Cowsik, R. & Wilson, L. *Proc. 14th int. Conf. Cosmic Rays*, Munich **2**, 659–664 (1975).

perturbation to the density field; and (3) a field of nearly adiabatic divergence-free ' ξ -motions' which ensure that the star remains in near-hydrostatic equilibrium despite the advection of the non-spherical pressure-density field by the nutation.

All quantities are given in terms of the spherical polar coordinate system (r, θ, λ) of the magnetic axis. In ref. 4 it is shown that the velocity field for the ξ -motions can be quantified by solving (1) $\nabla \cdot \xi = 0$; (2) $\xi_r = -\omega(\partial \rho_a / \partial \lambda) / (d\rho_a / dr)$ and (3) $\frac{1}{2} \int \rho_0 (\xi_\theta^2 + \xi_\lambda^2) dV = \text{minimum}$, subject to (1) and (2). ρ_0 is the density field of the unperturbed spherical star while ρ_a is the non-spherical part of the density perturbation. The divergence-free conditions for ξ follows directly from the adiabaticity assumption and will be enforced if the nutation period is less than the Kelvin–Helmholtz time, τ_{KH} . However, even if the nutation period is greater than τ_{KH} as outlined in ref. 4, this would mean that any ξ -motions with $\nabla \cdot \xi \neq 0$ would be subject to strong damping by radiative conduction. The condition $\nabla \cdot \xi = 0$ ensures that elements of gas move at constant density and suffer no compressional heating or cooling; to our order of working in the perturbation analysis, no temperature gradients develop other than those associated with hydrostatic equilibrium under the basic centrifugal and magnetic perturbations, and so there is no extra dissipation by radiative conduction. The minimizing condition (3) is used to remove the indeterminacy in the ξ -motion field. A complete analysis of the problem proves to be extremely difficult and this minimizing criterion should prove adequate for estimating astrophysically relevant numbers.

The ξ -field distorts the basic magnetic field B_0 so that the magnetic field may be written as $B_0 + \tilde{B}$ where $\tilde{B}$ is $O(\alpha^2 B_0)$ and it may be quantified by solving the perfect conductivity equation. The ξ -motions will also be subject to dissipation as the associated currents for $\tilde{B}$ will be subject to ohmic resistance. The dissipation of the ξ -motions drains the rotational energy of the star and this leads inevitably to a change in χ . For toroidal dominated fields, $\chi \rightarrow \pi/2$ and one is able to obtain an estimate for the time scale of perpendicularity, t_p , that is, the time for χ to increase from some initial angle less than $\pi/2$ to $\pi/2$. As we are assuming that solar-type stars have deep unobservable magnetic fields then the time scales for χ to reach its asymptotic state are only relevant in establishing the lifetime over which the ξ -motions will be present in the star.

The basic toroidal magnetic field was determined as the principal eigensolution of the decay equation with the boundary conditions that the field goes to zero at the centre of the star and that it does not penetrate the outer convection zone. The decay equation was solved with the conductivity proportional

to $T^{3/2}$ with T , the temperature, fixed by the temperature run for (1) and initial main-sequence solar model given in ref. 7 and (2) a solar model of Weymann⁸ for the present Sun.

In all of our calculations, we used these zero-order stellar models only as representative models for the Sun and solar-type stars with masses $\sim 1 M_{\odot}$.

In principle, for any particular model, the free parameters are the rotational velocity, α , the magnetic field strength, B_T (where B_0 is proportional to B_T), and the obliquity angle, χ . Thus, once a particular stellar model has been chosen and the relevant quantities have been computed, then extrapolations for the time scale of perpendicularity, the Eulerian nutation, the amplitude and speed of the ξ -motions can be carried out by making the relevant variations in α , B_T and χ . For example, t_p scales as α^2 , the time for one Eulerian nutation, $2\pi/\omega$, scales as $\alpha^{-1}B_T^{-2}$, while ξ scales as α^2 and $\dot{\xi}$ scales as $\alpha^2B_T^2$.

Looking at Table 1, for an initial angle of 75° if the star is rotating very rapidly (~ 0.2 days) then we have χ reaching its asymptotic state in $\sim 10^7$ yr which is within the present lifetime of the Sun for example. However, the results of Kraft⁹ indicate that after 3×10^7 yr, the solar surface rotation rate had decreased to ~ 5.2 days. Table 1 shows that as the rotation rate is decreased then t_p increases and for the slower rotation rates t_p is greater than the Sun's present lifetime of 4.6×10^9 yr. Thus Table 1 gives a quasi-evolutionary picture in that it indicates that as a star evolves and the rotation rate decreases, then the motions persist for longer and longer times. Consequently, for solar-type stars even if they commence their lives as rapid rotators, then we expect the ξ -motions to be present for most of their main sequence lifetime. Table 2 shows the time for one Eulerian nutation for $B_T = 10^5$ G for various rotation periods and χ values.

Table 3 gives estimates of the amplitude of the ξ -motions through an initial Sun in terms of the r -component of ξ . We have taken the rotation period to be 2 days as this is a value consistent with Kraft's findings that the solar surface rotation rate decreased from 0.2 to 5.2 days in the first $\sim 10^8$ yr. Also, we are given an estimate of the speed of the ξ -motions, again showing only the ξ_r component at various points through the star and the parameter B_T was taken to be 10^4 G. For comparison the meridian circulation velocity may be estimated to be 2.4×10^{-6} cm s $^{-1}$ so that if B_T is $> 10^4$ G then the ξ -motions have speeds substantially faster than the circulation speed. Actually a field of 10^4 G may be quite conservative since fields of 10^7 G (ref. 10) are possible as they would not be strong enough to penetrate through the convection zone and could thus be hidden without affecting hydrostatic equilibrium or thermonuclear processes. Thus, the entries for the speeds would then be multiplied by a factor of 10^6 so, for example, when $r \sim \frac{1}{2}R$ the speed is of the order of 3 cm s $^{-1}$. The computed amplitude of a particular oscillation is small ($\sim 4 \times 10^{-4} R$) for this rotation rate, however, if strong magnetic fields are present then a great deal of mixing is possible, if we invoke a random walk mechanism suggested by Mestel¹¹. A random walk or diffusion effect will occur if any particular fluid element does not return to its exact initial position due to dissipative processes such as radiative conduction and ohmic diffusion. If we take the mean square end-to-end distance as a measure of the extent of the random walk, then in time t the fluid particle could diffuse through a distance of the order of $\xi_r(t/t_{\text{nute}})^{1/2}$ where t_{nute} is the time for an Eulerian nutation. For the rotation period of 2 days and a field of $B_T \sim 10^4$ G we have from Table 3, $\xi_r \sim 2.6 \times 10^7$ cm at a point just below the convection zone. The values lead to a mixing length of 5.8×10^7 cm in the first 1.5×10^7 yr after a solar-type star commences its main sequence lifetime. This value is increased to 5.8×10^{10} cm or $\sim 0.8R$ if B_T is taken to be of the order of 10^7 G. If strong fields ($10^4 < B_T < 10^7$) are present a great deal of mixing is possible.

We expect the ξ -motions to be important in the early lifetime of solar-type stars because it is at this stage that the effects of rapid rotation and any strong magnetic fields present will combine to produce the large amplitudes. For Cowling-type fields

Table 1 Time scale [yr] of perpendicularity for toroidal field, initial Sun, $1 M_{\odot}$

χ	P	0.2 day	1.0 day	4.0 days
75°		1.48×10^7	3.69×10^8	5.91×10^9
45°		6.00×10^7	1.50×10^9	2.41×10^{10}
15°		4.92×10^8	1.23×10^{10}	1.96×10^{11}

Table 2 Time [yr] for one Eulerian nutation, $B_T = 10^5$ G, $1 M_{\odot}$ initial Sun

χ	P	0.2 day	1.0 day	4.0 days
75°		3.02×10^3	1.51×10^4	6.03×10^4
45°		1.10×10^3	5.50×10^3	2.20×10^4
15°		8.06×10^2	4.03×10^3	1.61×10^4

Table 3 Amplitude and speed of the ξ -motions for an initial Sun $1 M_{\odot}$

x	ξ_r (cm)	$\dot{\xi}_r$ (cm s $^{-1}$); $B_T = 10^4$ G
0.1	4.31×10^5	2.29×10^{-7}
0.2	9.75×10^5	5.18×10^{-7}
0.3	1.79×10^6	9.44×10^{-7}
0.5	5.58×10^6	2.96×10^{-6}
0.7	1.61×10^7	8.56×10^{-6}
0.8	2.61×10^7	1.38×10^{-5}

Rotation period is 2 days and $\chi = 75^\circ$; $x = r/R$.

B_0 increases inwards by a factor of ~ 25 . If other model magnetic fields were used in which B_0 increased by a significantly larger factor then we would expect (based on results in ref. 4) substantially more mixing through diffusion as the nutation frequency increases dramatically even for small values of B_T ($\sim 10^3$ G).

The recent results of Claverie *et al.*² indicate that the Sun may have a core which rotates 2–9 times more rapidly than the observed surface rotation. Thus for the Sun we have the possibility of the combination of a rapidly rotating core with a possible strong magnetic field operating over the solar lifetime, hence producing substantial diffusion throughout the Sun. For the period from 1.5×10^7 yr to the present the diffusion length would be a substantial fraction of the solar radius given the appropriate magnetic field structure. The extent to which this mixing can influence the solar neutrino anomaly is an open question.

Isaak¹² has associated the rotating core with the 12.2-day solar oblateness measurements of Dicke¹³. Furthermore, he has conjectured that the rapidly rotating core has a strong oblique magnetic field associated with it. We have calculated the nutation frequency for the Sun where we have taken the base of the convection zone to be at $r = 0.86R$. We assumed that the core is rotating with a period of 12.2 days and we find that the period of one nutation is $2.4 \times 10^{15}/B_T^2$, for an obliquity angle of 75° (or $8.9 \times 10^{15}/B_T^2$, for $\chi = 85^\circ$). If the nutation period is to correspond to the 22-yr solar cycle then $B_T = 1 \times 10^7$ G (or 2×10^7 G for $\chi = 86^\circ$).

Received 9 August; accepted 28 October 1982.

- Mestel, L. *Mem. Soc. R. Sci. Liège* **6**, 79–91 (1975).
- Claverie, A., Isaak, G. R., McLeod, C. P., van der Raay, H. B. & Roca Cortes, T. *Nature* **293**, 443–445 (1981).
- Schatzman, E. & Maeder, A. *Astr. Astrophys.* **96**, 1–16 (1981).
- Mestel, L., Nittmann, J., Wood, W. P. & Wright, G. A. E. *Mon. Not. R. astr. Soc.* **195**, 979–1000 (1981).
- Nittmann, J. & Wood, W. P. *Mon. Not. R. astr. Soc.* **196**, 491–506 (1981).
- Taylor, R. J. *Mon. Not. R. astr. Soc.* **191**, 151–163 (1980).
- Schwarzschild, M. *Structure and Evolution of the Stars* (Princeton University Press, 1958).
- Weymann, R. *Astrophys. J.* **126**, 208 (1957).
- Kraft, R. P. *Astrophys. J.* **150**, 551–570 (1967).
- Parker, E. N. *Cosmical Magnetic Fields* (Clarendon, Oxford, 1979).
- Mestel, L. in *Theoretical Principles in Astrophysics and Relativity* (eds Lebovitz, V. R., Reid, W. H. & Vandervoort, P. O.) (University of Chicago Press, 1978).
- Isaak, G. R. *Nature* **296**, 130–131 (1982).
- Dicke, R. H. *Sol. Phys.* **47**, 475–515 (1976).

Electric field effects in the thermal decomposition of solids

K. Rajeshwar, R. Rosenfold* & J. DuBow*

Department of Electrical Engineering, Colorado State University, Fort Collins, Colorado 80523, USA

The ability to 'tune' or even alter the thermal decomposition behaviour of solid materials is of practical interest, for example in their use as rocket propellants. During a study of the electrical conductivity of technologically-important materials^{1,2}, we became intrigued by the possibility of altering their thermal stability by imposition of an electric field across the heated samples. Our observations of d.c. electric field effects on the thermal decomposition of some model compounds are presented here. We have applied differential scanning calorimetry (DSC) to probe these effects. The DSC apparatus and the sample holder arrangement in particular was modified to permit application of an electric field across the heated sample. We have found that, in certain cases, the decomposition temperature can be lowered by as much as $\sim 100^\circ\text{C}$ in the presence of an electric field.

Electric field effects have been implicated in various thermophysical and thermochemical phenomena in solids and in solid-gas or liquid interfaces. It has been reported that an electric field accelerates the rate of growth of BaMoO_4 crystals in silica gel media³. A change in the sublimation pattern of KCl crystals has been observed in the presence of an electric field⁴. The lower dielectric strengths observed for ionic metal azides relative to other more thermally stable materials have been attributed to 'electrocatalytic' effects induced by the presence of highly conducting metal nuclei⁵. Selective phase formation at sample-electrode interfaces has been observed in a study of the effect of electric fields on the transformation of γ - to α - Al_2O_3 (ref. 6). An electric field enhancement of dehydroxylation rates in magnesium and aluminium hydroxides, has also been reported^{7,8}. The Russian literature also contains reports (see for example, ref. 9) of electric field effects on thermal decomposition of ionic solids (notably azides and oxalates). In these studies, however, a direct correlation between the applied field strength (E_{appl}) and the extent of modification of the thermal stability was not established. Such a correlation is afforded by our modified DSC technique in terms of monitoring any shift in the characteristic DSC peak temperature or a change in peak shape for a model reaction with E_{appl} (see below).

To permit imposition of an electric field across the heated sample in a commercial DSC apparatus (Dupont 990 Thermal Analysis System fitted with the DSC accessory module), the standard aluminium pans were replaced by matched rectangular blocks ($0.66 \times 0.60 \times 0.35$ cm) machined from Macor Ceramic (Corning Glass Co.). This material has manufacturer-quoted specifications (dielectric strength: 1.2×10^6 V cm^{-1} , thermal conductivity: 0.017 J s^{-1} cm^{-2} $^\circ\text{C}^{-1}$) which appear to be particularly suited for the present application. The sides and bottom of these modified sample and reference 'pans' had a nominal thickness (0.10 cm) which satisfied the conflicting requirements of efficient heat transfer and structural integrity. A comparison of peak shapes and transition temperatures for selected thermal analysis standards¹⁰ did not reveal significant differences between the conventional metal pans and ceramic material. Gold or platinum electrodes were press-fitted on opposite sides of the modified sample and reference 'pans'. External connections were made using insulated lead wires through holes drilled in the silver lid supplied with the DSC cell module. A d.c.

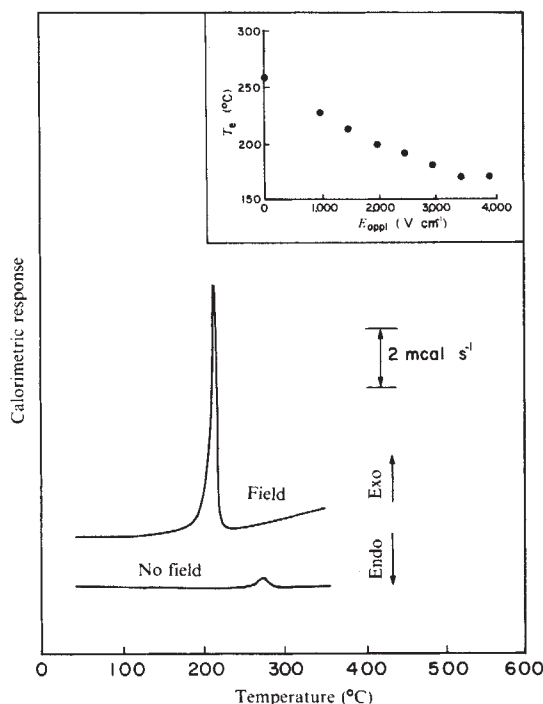


Fig. 1 Comparison of DSC traces for KMnO_4 in the presence and absence of an electric field ($2,000$ V cm^{-1}). The sample mass was 11.2 and 11.8 mg in the two cases respectively. Other experimental conditions are described in the text. The inset shows the dependence of the DSC peak extrapolated onset temperature, T_e , on the applied field, E_{appl} for KMnO_4 .

regulated power supply (Hewlett-Packard model 6522A) was used to apply test voltages in the range, 0 – $2,000$ V. The currents through the sample were monitored on a Keithley (model 616) electrometer.

All chemicals were of reagent-grade purity or better and were used as received. Each sample was ground to pass 200 mesh before subjecting it to DSC analyses. All experiments were done in a flowing atmosphere of dry N_2 (nominal flow rate: ~ 500 cm^3 min^{-1}) and at a heating rate of 10°C min^{-1} .

Figure 1 illustrates typical DSC curves for KMnO_4 in the presence and absence of a d.c. electric field ($2,000$ V cm^{-1}). Note the drastic increase in the exothermicity of the autooxidation peak and a lowering ($\sim 65^\circ\text{C}$) of the peak temperature for the sample exposed to the electric field relative to the control sample. (The results in this, and in later experiments, were not significantly altered by reversal of the d.c. electric field.) (In these experiments, as well as in all comparative runs, the same pair of matched pans was used for the control and test samples to preclude artefacts arising from variability in sample-holder properties.) The inset quantifies the relationship of the extrapolated onset temperature of the DSC peak (T_e) to E_{appl} . An average decrease in T_e of $\sim 0.03^\circ\text{C}$ for unit change in E_{appl} is observed for $E_{\text{appl}} < \sim 3,500$ V cm^{-1} .

Figure 2 shows representative DSC scans for NH_4ClO_4 , a commonly used composite rocket propellant oxidizer. The data are again shown for a control sample and a specimen exposed to a d.c. electric field ($1,000$ V cm^{-1}). A shift in T_e for the high-temperature autooxidation peak (marked D) of $\sim 25^\circ\text{C}$ and an increase in its exothermicity are noted for the sample exposed to the electric field relative to the control case. The intermediate temperature decomposition peak (marked C) and the endotherm corresponding to the orthorhombic $\rightleftharpoons$ cubic transformation (marked B) are less sensitive to E_{appl} . Note, however, the pronounced appearance of a low-temperature decomposition exotherm (marked A) when an electric field is imposed. This peak is absent in the control sample—behaviour which is expected from the very long induction periods required for this low-temperature step¹¹.

* Present address: College of Engineering, Boston University, Boston, Massachusetts 02215, USA.

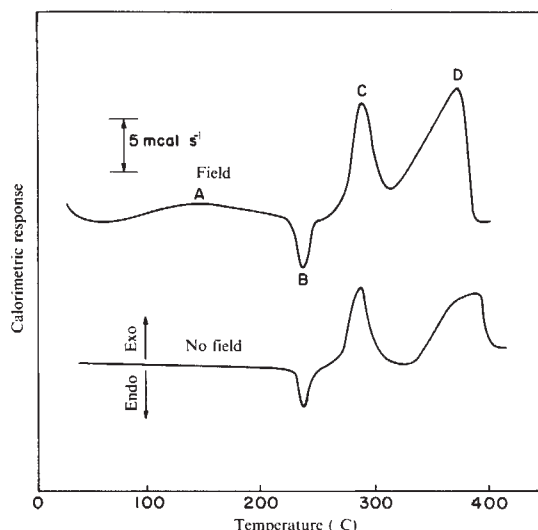


Fig. 2 Comparison of DSC traces for NH_4ClO_4 in the presence and absence of an electric field ($1,000 \text{ V cm}^{-1}$). The sample mass was 18.3 mg and 18.2 mg in the two cases respectively. Other experimental conditions as well as assignment of the peaks marked A–D may be found in the text.

It is obvious from the results presented above and from our cursory examination of other compounds and materials (KClO_4 , $\text{CaC}_2\text{O}_4 \cdot \text{H}_2\text{O}$, $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, NaN_3 , oil shale and coal), that electric field effects manifest themselves in various ways depending on the specific nature of the thermal reaction and the material under investigation. For example, our experiments on NaN_3 in the presence of an electric field reveal a pronounced increase in exothermicity paralleling the behaviour shown in Fig. 1 for KMnO_4 . The shift in T_c , however, is much lower than in the KMnO_4 case. Consistent with the results reported previously on this material⁵, we have also observed current instabilities and dielectric breakdown in the temperature region (350 – 400°C) immediately following the thermal decomposition peak.

While it would be hazardous to propose a single mechanism to account for electric field effects in general, some preliminary conclusions may be drawn on the basis of the present data. The efficacy of E_{appl} in lowering T_c seems to be related to the extent to which charge transfer plays a rate-determining role in the very early stages of the decomposition. To this end, the d.c. electric field is expected to assist in this charge transfer step (as, for example, in the hopping of an electron from one anion site to another which has been invoked for KMnO_4 ; that is, $\text{MnO}_4^- + \text{MnO}_4^- \rightarrow \text{MnO}_4^0 + \text{MnO}_4^{2-}$, see ref. 12). Similarly, electron or proton transfer is believed to play a key role in the low-temperature thermal decomposition of NH_4ClO_4 (refs 11, 13). The enhancement of this reaction in the presence of a d.c. electric field (which would assist the elementary charge transfer in either case) (see Fig. 2) would be consistent with this belief. Another factor in these effects (one which, in our opinion, is not readily estimated) is the extent of localized self-heating in the material and the concomitant creation of 'hot spots' which can explain the observed shifts in T_c and the increase in exothermicity when an electric field is present. (In this respect, very high electric field strengths are to be expected at grain-grain interfaces within the bulk sample especially because of the intrinsically low bulk conductivity of the material and the average thickness, 10^{-4} – 10^{-6} cm , of depletion layers typical of ionic solids, see ref. 9.) It is unlikely, however, that Joule heating of the bulk sample can account for the observed effects. (A gradual upward shift in the DSC baseline which is often observed at temperatures $\gg T_c$ (see Fig. 1), however, is undoubtedly a manifestation of Joule heating.) The current flow through the test sample was always less than the lower limit of sensitivity of the measurement set-up ($\sim 10^{-8} \text{ A}$) except in cases where dielectric breakdown occurred (see above). A rough calculation

of the temperature increase due to Joule heating (assuming specific heat values typical of the materials studied here, $\sim 1.3 \text{ J g}^{-1}^\circ\text{C}^{-1}$, ref. 14), reveals levels which are orders-of-magnitude lower than the T_c shifts observed in this study (see Figs 1 and 2). (These calculations assume that the entire E_{appl} is dropped across the sample which is an overestimation.)

We thank Tom Mraz for useful discussions.

Received 16 August; accepted 21 October 1982.

1. Rajeshwar, K., Das, M. & DuBow, J. *Nature* **287**, 131–133 (1980).
2. Rajeshwar, K., Nottenburg, R. Pai, Verneker, V., & DuBow, J. *J. chem. Phys.* **72**, 6678–6684 (1980).
3. Kurien, K. V. & Ittyachen, M. A. *J. mater. Sci.* **15**, 1724–1729 (1980).
4. Jose Yacamán, M., Munir, Z. A., Ocaña, T. & Hirth, J. P. *Appl. Phys. Lett.* **34**, 727–728 (1979).
5. Tang, T. B. & Chaudhri, M. M. *Nature* **282**, 54–55 (1979).
6. MacKenzie, K. J. D. & Hosseini, G. *Trans. Br. ceram. Soc.* **75**, 172–176 (1976).
7. MacKenzie, K. J. D. *Thermochim. Acta* **25**, 277–288 (1978).
8. MacKenzie, K. J. D. *Thermochim. Acta* **16**, 1–16 (1976).
9. Kabanov, A. A. *Russ. Chem. Rev.* **40**, 953–963 (1971).
10. McAdie, H. G. *Thermochim. Acta* **1**, 325–333 (1970).
11. Jacobs, P. W. N. & Whitehead, H. M. *Chem. Rev.* **69**, 551–590 (1969).
12. Boldyrev, V. V. *J. Phys. Chem. Solids* **30**, 1215–1223 (1969).
13. Khairatdinov, E. F. & Boldyrev, V. V. *Thermochim. Acta* **41**, 63–86 (1980).
14. Rajeshwar, K., Verneker, V. Pai, & DuBow, J. *Combust. Flame* **37**, 251–260 (1980).

Tension fractures and extensional tectonics

Jacques Angelier

Département de Géotectonique, Université Pierre-et-Marie Curie, 4, place Jussieu 75230 Paris, Cédex 05, France

Bernard Colletta

Instituto de Geología, Universidad Nacional Autónoma de México, Hermosillo, Sonora, México

A characteristic pattern of normal faults that we have studied in several regions illustrates successive geometries relating to increasing amounts of extension from 10–50% (Gulf of Suez, Egypt) to 50–100% (Western Gulf of California, Mexico) up to 200% (Southern Basin and Range, USA). Tension fractures formed at an early stage of extension have an essential role. The resulting structure is described as a multiple tilted 'packs of cards' model in which rotational deformation within the major blocks is accommodated by normal faulting on previously formed tension fractures, thus defining smaller blocks of second and third generation.

Where amounts of extension remain as small as 10–20% ($\beta = 1.1$ – 1.2), more or less parallel normal faults separate elongated blocks (Fig. 1a, b). The blocks are gently tilted (0° – 10°) and usually 5–15 km wide. The dips of the faults observed in the field commonly average 60° – 65° (the dip of faults at depth is probably significantly smaller than close to the surface and averages 45° – 50°). Many vertical (80° – 90°) tension gashes and fractures striking parallel to normal fault trends develop during this first stage of extension, either when the block tilting begins or just after (5° – 10°). This results in a dense pattern of fractures that are perpendicular or approximately perpendicular to layering, whatever the further tilting. An example of this stage of extension is provided by the Esch El Melaha range, on the western edge of the Gulf of Suez rift. Similar structural patterns are common in continental regions that have been submitted to moderate extension, As Aegea¹.

In the next stage of increasing extension, with amounts ranging from 20% ($\beta = 1.2$) to 70% ($\beta = 1.7$), pure dip-slip motion continues along main normal faults and block tilting increases from 10° to $\sim 25^\circ$, thus inducing the typical 'pack-of-cards' structure² (Fig. 1c). Provided the average elevation of blocks is the same, the amount of extension is given by a simple formula where ζ_0 and θ design the initial dip of the faults and the angle

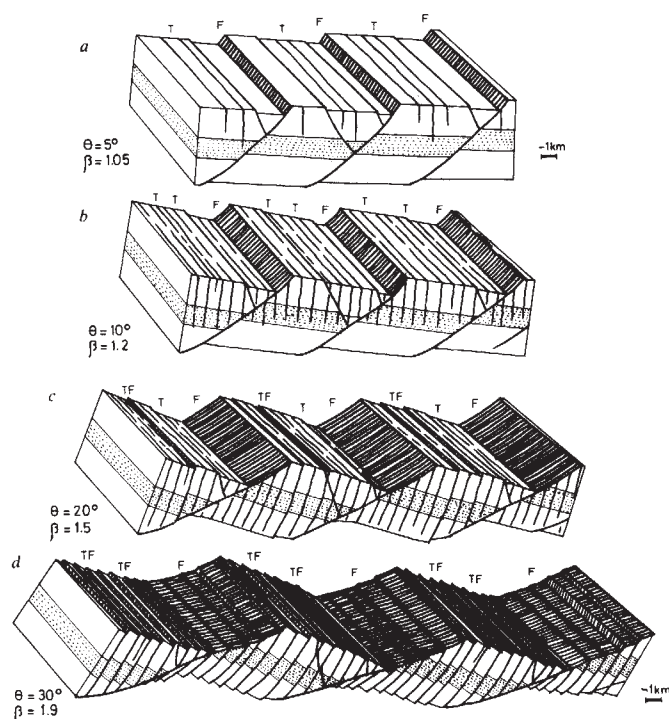


Fig. 1 Evolution of faulting with increasing amounts of extension (a–d). F, normal faults oblique to bedding (especially first-order faults). T, tension gashes and fractures. TF, normal faults approximately perpendicular to bedding (second- and third-order faults).

of tilt, respectively:

$$\beta = \frac{\sin \zeta_0}{\sin (\zeta_0 - \theta)} \quad (\theta < \zeta_0).$$

As motion concentrates on large normal faults, tension gashes and fractures are simply tilted together with the strata to which they are perpendicular. When the tilt approaches 25°–30°, a value which commonly corresponds to 50–150% of extension ($\beta = 1.5$ –2.5, depending on the value of ζ_0), a new phenomenon occurs and leads to the development of the type of faults we shall focus on: some of the tension fractures begin to act as normal faults (Fig. 1c, d). Both the throws of these incipient faults and their number relative to that of earlier fractures increase with tilting and extension. This critical stage has been encountered in the Gebel Zeit range, on the western shore of the Gulf of Suez: layers which commonly dip 35° to the west are cut by very closely spaced fractures perpendicular to bedding. Most of these fractures do not offset the beds, but some have begun to act as dip-slip normal faults and show significant offsets (of centimetre to metre size).

In the western portion of the Zeit block, the stratal tilting reaches 40°–50°, while the number and amplitude of throws along faults perpendicular to bedding increases. We infer that a strong relationship exists between stratal tilting and activation of tension gashes and fractures as normal faults, as extension

increases. This phenomenon leads to the next stage of extension, which can be illustrated by the Arroyo Inferno section of the Santa Rosalia basin (Baja California, Mexico) (Fig. 2): most faults make angles of 80°–90° with bedding planes that dip 40°–45° to the east. These normal faults perpendicular to layering account for a large amount of extension, although faults that were originally dipping 50°–70° are also present. Almost all fractures perpendicular to bedding have moved as normal faults. With $\zeta_0 \approx 90^\circ$, the coefficient of extension related to motion along faults perpendicular to bedding can be computed using the formula already mentioned. One thus obtains a minimum amount of 40% ($\beta = 1.4$) for the Arroyo Inferno section, knowing that other types of normal faults induce further extension.

Similar fault patterns can be observed in various sections of the Basin and Range province of the western USA. The average angles between these faults and the bedding planes are commonly slightly smaller than 90° (that is, 80°–85°); as a result, most of these faults would appear as steep reverse faults after rotating the strata back to their initial horizontal position. This observation suggests that most vertical tension gashes and fractures have actually developed just after the beginning of extensional processes, when the tilt of the main blocks was as moderate as 5°–10°.

Increasing stratal tilting and larger amounts of extension can be observed in the Basin and Range province³ and especially in the Eldorado Mountains, south of Las Vegas (Nevada), where highly deformed Cenozoic formations have been described⁴. Faults perpendicular to bedding are abundant with two orders of magnitude. Relatively continuous and large faults, with throws of hundred metres, are spaced about 0.8–1 km apart. Smaller but numerous faults, with throws of centimetres to metres, can be observed in detail to have a much closer spacing (decimetres to decametres). Both types of faults strike parallel to the trend of tilted layers and make angles of 80°–90° with strata. We call these second- and third-order faults, whereas first-order normal faults define major blocks that commonly are several kilometres wide. In addition, the first-order normal faults have a different attitude relative to bedding (angles of 45°–70° instead of 80°–90°). The second and third order faults correspond to a severe internal deformation of the major first-order blocks which are divided into numerous second- and third-order blocks and slices (Fig. 3). In detail, the stratal tilting may vary from 30° to 90° within a single first-order block, thus showing that the deformation is complex and implies differential rotations (Fig. 3b).

This model is primarily characterized by associated normal faulting and stratal tilting which define sets of parallel elongated blocks separated by planar dip-slip normal faults. The difference from the single 'pack-of-cards' model^{2,5} is the presence of second- and third-order normal faults perpendicular to bedding. Such faults account for a large amount of extension provided the total extension exceeds 50–100% ($\beta = 1.5$ –2): these values usually correspond to tilt angles of 25°–35°. Motion along these faults probably develops when motion along large first-order faults becomes difficult as the result of flattening related to progressive tilting. This multiple 'pack-of-cards' model resembles the pattern of normal faulting already described in

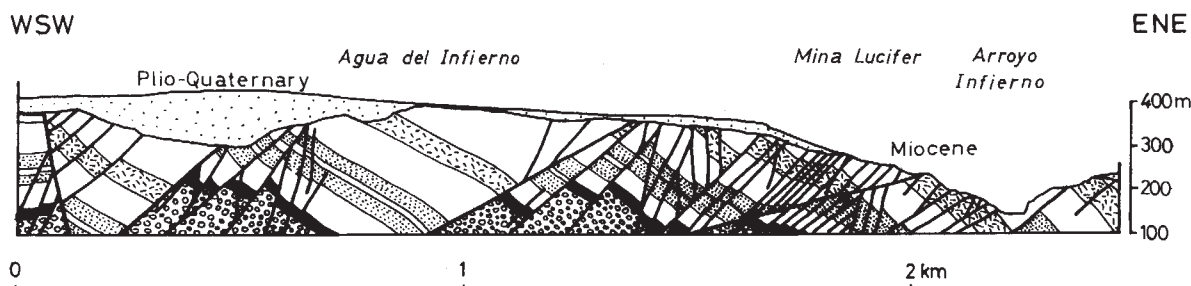


Fig. 2 Arroyo Inferno section, northwestern Mexico.

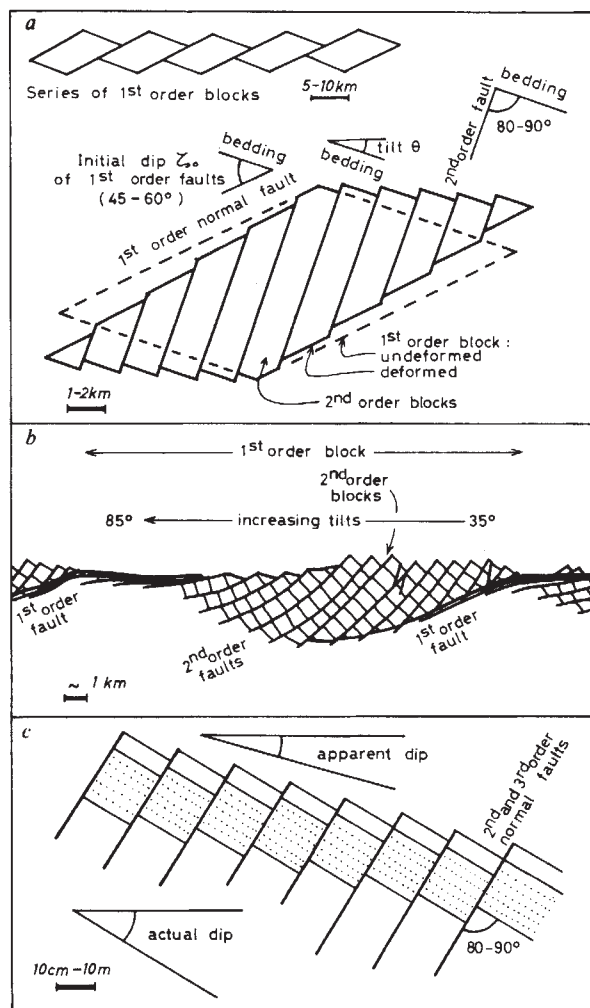


Fig. 3 Internal deformation of faulted-tilted blocks. Schematic (a) and actual (b) structures; c, thinning and flattening of layers.

the Danakil rift of Ethiopia⁶. We can explain the presence of second- and third-order faults by intense fracturing that creates numerous narrowly spaced tension gashes and fractures during the earliest stages of extensional deformation⁷. The pattern of deformation related to further motion along these second- and third-order faults results in internal extension and rotation within the first-order faulted-tilted blocks (Figs 1, 3). This internal formation does not have a significant role as long as the total amount of extension remains smaller than 50-100%. Consequently, quantifying extension versus tilting requires the use of a composite function. The first function corresponds to the single 'pack-of-cards' model²; the second one also takes into account the internal deformation within the major blocks. This component of extension is difficult to quantify in a general way because second and third order blocks may progressively rotate across the main blocks (Fig. 3b). Our estimates suggest that, based on field examples, average tilts of 20°, 35°, 50° and 60° approximately correspond to 50, 100, 150 and 200% of extension, respectively ($\beta = 1.5, 2, 2.5$ and 3).

Finally, the single 'pack-of-cards' model probably remains valid until the stratal tilting reaches values of $\sim 25^\circ$: the corresponding amount of extension averages 100% ($\beta = 2$) and the dip of first-order faults is 20°, provided the original dip was $\zeta_0 = 45^\circ$. Then, motion along first-order normal faults is more and more difficult so that these main faults will progressively become inactive for increasing stratal tilting (from 20° to 40°). As a result, extensional processes now concentrate on internal deformation of main blocks, that occurs along numerous, gradually activated second- and third-order faults. Finally, after

this transitional period, the whole of the extension is accounted for by relative motions of second- and third-order blocks.

Neglecting one type of faulting (such as faults perpendicular to bedding) can lead to underestimates of the total amount of extension, especially in strongly deformed areas. Other sources of underestimation are present: neglecting small-scale extensional structures and continuous deformation of upper faulted layers, or the effects of sedimentation during tilting. What may be, for example, the influence of internal deformation of the major blocks in the structure of passive margins as the Armorican margin^{8,9}? There is no reason to assume that such internal deformation is absent where block tilting is important, based on land observations discussed above. Depending on the density of second- and third-order faulting, this deformation may result either in an apparent thinning and flattening of layers reconstructed on seismic reflection profiles (Fig. 3c), or even in obscuring the internal structure of the blocks; both effects leading to underestimation of block deformation and consequently of the amount of extension.

The multiple 'pack-of-cards' model thus enables one to understand how high amounts of extension can be present in passive margins where different estimates have been made using interpretations of the same seismic reflection profiles^{8,9}. Our field observations thus indicate that for large amounts of extension the first-order blocks are affected by sizeable rotational deformation at the scale of hundreds of metres and metres (second- and third-order blocks). Consequently, balanced cross-sections based on seismic reflection profiles which do not detect small scale deformation, do not restore the actual pre-deformation geometry.

Received 30 July; accepted 25 October 1982.

1. Angelier, J. *Publ. Soc. géol. Nord* **3** (1979).
2. Le Pichon, X. & Sibuet, J. C. *J. geophys. Res.* **86**, B5, 3708-3720 (1981).
3. Zoback, M. L., Anderson, R. E. & Thompson, G. A. *Phil. Trans. R. Soc. A* **300**, 407-434 (1981).
4. Anderson, R. E. *Bull. geol. Soc. Am.* **82**, 43-58 (1971).
5. Wernicke, B. & Burchfiel, B. C. *J. struct. Geol.* (in the press).
6. Morton, W. H. & Black, R. in *Afar Depression of Ethiopia* Vol. 1 (eds Pilger, A. & Rösler, A. 1, 55-65 (Schweizerbart'sche, Stuttgart, 1975).
7. Colletta, B. & Angelier, J. *C. r. heb. Séanc. Acad. Sci., Paris B294*, 467-469 (1982).
8. Montadert, L., Chenet, P. Y., Gairaud, H. & Roberts, D. G. *Hedberg Res. Conf. Continental Margin Processes*, Galveston (1980).
9. Le Pichon, X., Sibuet, J. C. & Angelier, J. *Proc. Am. Ass. petr. Geol.* (in the press).

The Cerro Galan ignimbrite

P. W. Francis*, L. O'Callaghan*, G. A. Kretzschmar*, R. S. Thorpe*, R. S. J. Sparks†, R. N. Page‡, R. E. de Barrio‡, G. Gillou§, & O. E. Gonzalez§

* Department of Earth Sciences, The Open University, Milton Keynes MK7 6AA, UK

† Department of Earth Sciences, University of Cambridge, Downing Place, Cambridge CB2 3EQ, UK

‡ Servicio Geológico Nacional, Av. Santa Fe 1548, 1060 Buenos Aires, Argentina

§ Servicio Minero Nacional, Miguel Lillo 205, 4000 San Miguel de Tucuman, Argentina

The Cerro Galan caldera, north-west Argentina, is one of the world's largest resurgent calderas but was not recognized as such until the advent of spacecraft imagery^{1,2}. Here we present some initial results of the first scientific expedition to the caldera that was formed during a major eruption 2.5 Myr ago during which 1,000 km³ of ignimbrite was erupted. The pre-welding thickness of ignimbrite within the resurgent part of the caldera was ~ 2 km. The ignimbrite is fine grained and lacks large pumice and lithic clasts. An underlying plinian air fall pumice deposit is absent. The wide dispersal of the ignimbrite suggests an unusually violent eruption, perhaps initiated by subsidence of the caldera floor into the magma chamber.

The Cerro Galan caldera is an elliptical structure 34 × 20 km, with its longest axis oriented north-south, parallel to regional structural trends (Fig. 1). The western rim exceeds 5,200 m,

Table 1 K-Ar determinations of Cerro Galan caldera ignimbrites

Specimen	Locality	Mineral	K (triplicate) (%)	Radiogenic argon (%)	Volume ^{40}Ar radiogenic Sec 1g	Age and error (Myr)
A78-1	25°47' S	Biotite	7.537 ± 0.24	19.87	7.301×10^{-7}	2.49 ± 0.12
Cerro Galan ignimbrite	67°18' W					
A78-16	26°00' S	Biotite	7.713 ± 0.052	31.3	7.393×10^{-7}	2.46 ± 0.12
Cerro Galan ignimbrite	67°16' W					
A78-14	26°02' S	Biotite	7.489 ± 0.21	11.06	1.420×10^{-6}	4.87 ± 0.39
Real Grande ignimbrite	67°25' W					
A78-51	25°48' S	Biotite	6.454 ± 0.16	18.53	1.206×10^{-6}	4.80 ± 0.24
Real Grande ignimbrite	67°15' W					
A78-20	26°08' S	Biotite	6.518 ± 0.19	17.75	1.365×10^{-6}	5.38 ± 0.26
Older ignimbrite	67°18' W					

K-Ar determinations by G.A.K. at the University of Leeds. $\lambda_e = 0.581 \times 10^{-10} \text{ yr}^{-1}$, $\lambda_\beta = 4.962 \times 10^{-10} \text{ yr}^{-1}$, $^{40}\text{K} = 1.167 \times 10^{-2} \text{ atom } \%$. Decay constants from ref. 18. Analytical errors refer to the s.e.m. at the 2σ level.

the caldera floor is at a mean height of 4,600 m and the resurgent centre reaches ~6,100 m at its highest point. The ignimbrite erupted from the caldera is conspicuous on Landsat imagery, forming an apron which can at present be traced continuously for distances up to 70 km, and undoubtedly extended further originally. The Cerro Galan ignimbrite overlies a major sequence of earlier ignimbrites locally totalling 300 m thickness. K-Ar age determinations (Table 1) show that the age of the Cerro Galan ignimbrite is 2.5 Myr and that of the youngest dated underlying flow is 4.8 Myr. A well-defined erosional break is present between the two.

The outflow sheet of the Cerro Galan ignimbrite is spectacularly exposed on all flanks of the caldera except the south. In the west, it forms a single sheet 20–30 m thick on flat surfaces, but up to 200 m thick where it fills valleys cut in the older underlying ignimbrites. Only one flow unit is usually present. In the north, the ignimbrite is uniformly substantially thicker, is over 100 m thick 20 km from the caldera rim and consists of up to four flow units. Although our data are insufficient to construct a complete isopach map, the ignimbrite appears to be much the same thickness 70 km from the caldera as it is 30 km from it, indicating a relatively low aspect ratio³. On the flanks of the caldera the ignimbrite is typically an incipiently welded 'sillar' forming prominent cliff lines and locally exhibiting well developed columnar jointing where it entered pre-existing river valleys, indicating emplacement at high temperature. On the northern flanks, where it is thickest, the ignimbrite is densely welded. Pumice clasts are small and scarce; those that are present rarely exceed 2–3 cm and are fragile and highly vesiculated. Lithic clasts are also small and scarce, rarely exceeding 1–2 cm even in proximal facies, except in the north where maximum clast sizes are 2–3 cm. Breccia lenses in the immediate vicinity of the rim contain clasts up to 30 cm. These may be co-ignimbrite lag breccias⁴.

Analyses of separate pumice clasts show that the magmatic composition was dacitic ($\text{SiO}_2 = 68\%$). The bulk of the ignimbrite is made up of crystals of plagioclase (An_{33}), biotite, quartz, sanidine and magnetite in a matrix of fine pumice shards and dust. Model analyses show that typical ignimbrite samples contain about 55% by volume crystals, whereas pumice clasts contain only ~35%, indicating that crystal concentration took place by elutriation of fine ash particles^{5,6}, and that up to 30% of the original volume of the ignimbrite may have been dispersed.

Although the base of the ignimbrite is very well exposed all around the flanks of the caldera, there is no evidence anywhere of a pre-ignimbrite plinian air fall deposit, even on the eastern flanks, downwind of the prevailing winds. In this respect, the Cerro Galan ignimbrite is markedly different from many well documented ignimbrites. Evidence that the eruption was a violent single catastrophic event is provided not only by the wide distribution of the ignimbrite but also by several sedimentary structures. Flame structures at its base show that the ignimbrite was strongly erosive, and dispersion of locally

derived lithics throughout its thickness indicates turbulent flow conditions. The flow also climbed several hundred metres up the slopes of the Cerro Beltran on the western side of the caldera (Fig. 1), suggesting high velocities and rapid emplacement^{7,8}.

Within the caldera itself, the central resurgent massif consists entirely of ignimbrite, which is partially to densely welded throughout, except in the extreme uppermost exposed levels. We are confident on petrographic grounds that this ignimbrite is the same as that exposed on the flanks of the caldera. The present thickness of ignimbrite within the caldera is at least 1,200 m and a pre-welding thickness of the order of 2,000 m is likely. The eastern margin of the resurgent massif is bounded by a major north-south fault. Although ignimbrites on the uplifted block are juxtaposed against basement rocks along the fault, the base of the ignimbrite is not exposed, confirming the great thickness of the ignimbrite within the caldera.

In assessing the total volume of the eruption, three factors must be taken into account: the volumes of the intra-caldera ignimbrite, the outflow sheets and the dispersed, fine grained 'co-ignimbrite ash'. Given that the area of the caldera floor is $\approx 450 \text{ km}^2$, the minimum volume of ignimbrite within the caldera is of the order of 500 km^3 . Typical measured densities of these densely welded rocks range between 2.1 and $2.4 \times 10^3 \text{ kg m}^{-3}$, indicating a dense rock equivalent volume of $\approx 450 \text{ km}^3$. At present, ashflows on the caldera flanks are exposed over an area of some $7,500 \text{ km}^2$, with a volume of $\sim 225 \text{ km}^3$ at the most conservative estimate. Measured densities range from 1.9 to $2.3 \times 10^3 \text{ kg m}^{-3}$ for rocks showing different degrees of welding. A dense rock equivalent volume of 175 km^3 is indicated, assuming an average density of $2.0 \times 10^3 \text{ kg m}^{-3}$. The pre-erosion volume is harder to deduce; it may have been double that at present.

K-Ar dating of ash horizons in fossil bearing late Tertiary continental sediments from several localities in north-west Argentina suggests that some of these are probably derived from Cerro Galan from pre-caldera eruptions⁹ and from the main caldera forming event itself¹⁰. In particular, the important 2.78 ± 0.09 -Myr old 'toba dacitica' at Uquia, 270 km north-east of Cerro Galan may correlate with the Cerro Galan ignimbrite¹⁰, given the uncertainties in radiometric determinations and possibility of contamination by detrital material, and that other major source structures are not obvious. Such correlations remain to be confirmed and extended, but it is to be expected that co-ignimbrite ashes will be found over large areas of north-west Argentina, and these can only add to volume estimates for the Cerro Galan eruption. Co-ignimbrite ashes often have volumes comparable with the associated ignimbrites^{5,6}, consistent with the crystal concentration actually observed at Cerro Galan. Thus, a total volume for the eruption well in excess of $1,000 \text{ km}^3$ is indicated. Clearly the eruption must rate as one of the largest recent examples in the geological record.

The absence of a pre-ignimbrite plinian air fall deposit, the widespread dispersal of the ignimbrite and the scarcity within



Fig. 1 A Landsat image of the Cerro Galan caldera. The snow-capped, 6,100 m resurgent centre is conspicuous at the centre. Laguna Diamante is a relic of a much larger lake which may have filled caldera before resurgence. Extensive apron of Cerro Galan ignimbrite is discernible in pale-grey tones, deeply gullied terrain, best developed in north and west flanks. Cerro Colorado and Cerro Beltran are pre-caldera andesite volcanoes. Dark areas west of Cerro Colorado are basaltic andesite scoria cones and lava flows broadly contemporaneous with caldera forming events.

it of both pumice and lithic clasts suggest that the Cerro Galan ignimbrite was the result of a rather different eruption mechanism than that proposed for other well-described ignimbrites such as the Bandelier and Bishop Tuffs^{11,12}. The eruption rate may have been unusually rapid.

The fact that the Cerro Galan ignimbrites were able to travel horizontal distances in excess of 70 km indicates that they acquired much of their kinetic energy by collapse from a vertical eruption column, probably at least 10 km high⁷. A sustained, convecting vertical eruption column, however, would have led to the deposition of a recognizable plinian air fall deposit. This could suggest that the radius of the erupting vent widened almost instantaneously to the radius at which a convecting column would be unstable¹³. Alternatively, differences in magma volatile content may have been involved. Large volume silicic eruptions are now conceived as beginning with the eruption of a volatile-rich cap of the magma chamber, which leads to the formation of a crystal poor plinian deposit, followed by ignimbrite units and ultimately by extrusion of lavas¹⁴.

Previously models for resurgent calderas suggest that eruption commenced along inward dipping ring fractures¹⁵. Although there is no direct field evidence one way or the other, in the case of Cerro Galan, we suggest that it is more likely that the ring fractures were outward dipping, allowing for the catastrophic foundering of the caldera floor, much as originally suggested by Bailey *et al.* for the Glen Coe cauldron¹⁶. Such a triggering mechanism requires the magma chamber to be under-

pressured and the outward dipping shear fractures to be generated when the lithostatic pressure on the roof of the magma chamber exceeds the chamber pressure by the compressive strength of the rock¹⁷. Very large eruption rates (10^8 – 10^9 m³ s⁻¹) can be generated as the dense caldera floor block sinks into the lower density silicic magma, provided the ring fractures are outward dipping (T. H. Druitt and R.S.J.S. in preparation). Inward dips would inhibit such rapid eruption rates. The mechanism triggering the subsidence of the caldera floor into the underlying magma chamber, however, remains speculative.

We emphasize the preliminary nature of the conclusions presented here. Petrological studies are in progress and will be published elsewhere, but many years work will be required to document fully this large structure.

Fieldwork in Argentina was financed by the Royal Society, the Servicio Minero Nacional and Servicio Geologico Nacional, the Royal Geographical Society, the Mount Everest Foundation and other bodies. K-Ar age determinations were made by G.A.K. during the tenure of a NERC studentship. Drs Mario Alderete and Roberto Caminos are thanked for logistic support in Argentina from Servicio Minero Nacional and Servicio Geologico Nacional.

Received 20 July; accepted 26 October 1982.

1. Friedman, J. D. & Heiken, G. in *Skylab Explores the Earth*, 380 (NASA Spec. Publ., 1977).
2. Francis, P. W., Hammill, M., Kretzschmar, G. A. & Thorpe, R. S. *Nature* **274**, 748–751 (1978).
3. Walker, G. P. L., Heming, R. F. & Wilson, C. J. N. *Nature* **283**, 286–287 (1980).
4. Druitt, T. H. & Sparks, R. S. J. *J. Volcan. Geotherm. Res.* **13**, 147–171 (1982).
5. Walker, G. P. L. *Contr. Miner. Petrol.* **36**, 135–146 (1972).
6. Sparks, R. S. J. & Walker, G. P. L. *J. Volcan. Geotherm. Res.* **2**, 329–34, (1977).
7. Francis, P. W. & Baker, M. C. W. *Nature* **270**, 164–165 (1977).
8. Miller, T. P. & Smith, R. L. *Geology* **5**, 173–176 (1977).
9. Marshall, L. G., Butler, R. F., Drake, R. E., Curtiss, G. H. & Tedford, R. H. *Science* **206**, 272–278 (1979).
10. Marshall, L. G., Butler, R. R., Drake, R. E. & Curtiss, G. H. *Science* **216**, 272–278 (1978).
11. Smith, R. L. & Bailey, R. A. *Bull. volcan.* **29**, 83–104 (1966).
12. Hildreth, W. *Geol. Soc. Am. Spec. Pap.* **180**, 43–76 (1979).
13. Wilson, L., Sparks, R. S. J. & Walker, G. P. L. *Geophys. J. R. astr. Soc.* **63**, 117–148 (1980).
14. Hildreth, W. *J. geophys. Res.* **86**, 10153–10192 (1981).
15. Smith, R. L. *Geol. Soc. Am. Mem.* **116**, 613–622 (1968).
16. Bailey, E. B. & Maufe, H. B. *The Geology of Ben Nevis and Glen Coe and the Surrounding Country*, 2nd edn (Geological Survey of Scotland, 1980).
17. Roberts, J. L. *Geol. J. Spec. Issue* **2**, 287–338 (1970).
18. Steger, R. H. & Jäger, E. *Earth planet. Sci. Lett.* **36**, 359–362 (1977).

Effects of increased CO₂ concentrations on surface temperature of the early Earth

William R. Kuhn

Department of Atmospheric and Oceanic Science, The University of Michigan, Ann Arbor, Michigan 48109, USA

James F. Kasting

Ames Research Center, Mountain View, California 94035, USA

One of the major riddles of terrestrial evolution concerns the history of the Earth's surface temperature. Present models of solar evolution indicate that the luminosity was about 25% smaller at the time of formation of the Solar System¹. This factor alone would have caused the Earth's surface temperature to be below freezing, yet the geological record indicates that water has flowed on the surface since at least 3,800 Myr ago². Enhanced levels of CO₂ in the atmosphere could have provided the necessary warming to maintain the temperature above freezing. Increased tectonic activity, and a decrease in solubility of CO₂ in the oceans, rock weathering, and sediment deposition are processes that have been suggested for these larger amounts of CO₂. We show here that large CO₂ concentrations are necessary to maintain the early Earth's surface temperature at approximately today's level. If there were a thousand times the present atmospheric level (PAL) of CO₂ in the atmosphere, the temperature would be 292K, while a 100-fold increase in CO₂ concentration yields 284K. The surface warming is highly dependent on the amount of water vapour and clouds, both of which we have little knowledge of for this very early time in the Earth's history.

A likely mechanism for warming the early Earth was the presence of one or more greenhouse gases in the atmosphere in addition to water vapour. Sagan and Mullen³ suggested that ammonia could have provided the necessary heating if present in concentrations of a few parts per million. However, even this small mixing ratio would have been hard to support due to the photochemical instability of NH_3 with respect to conversion to N_2 (refs 4, 5). Carbon dioxide, which should have been one of the major constituents of Archaean volcanic emissions, is more likely^{6,7}. The ability of CO_2 to provide an effective greenhouse has been tested by Owen *et al.*⁸. Drawing on the evolutionary sequence proposed by Hart⁹, in which solar luminosity has increased by 25% and the partial pressure of CO_2 has decreased exponentially with time from an initial value of 0.31 bar, Owen *et al.*⁸ predict that surface temperatures during the Earth's early history could have been 10–20 K higher than at present.

While these results do suggest that CO_2 could have provided the required warming, it is unwise to interpret them literally as a time history of the Earth's surface temperature. Hart's evolutionary model provides at best only a crude parameterization of the factors controlling the atmospheric CO_2 partial pressure. Hart's basic assumption that the rate of CO_2 outgassing is balanced by the consumption of CO_2 due to weathering of silicate minerals and subsequent deposition of carbonate sediments is in accord with present theories concerning the carbon cycle^{7,10}. He assumes that the degassing rates for juvenile CO_2 and other constituents have declined exponentially with a time constant of 800 Myr, similar to the time constant of 862 Myr estimated by Li¹¹ for juvenile degassing processes. Recycling of CO_2 was not included. The latter process, which accounts for almost all the CO_2 outgassed today, occurs at an estimated global rate of $10^{14} \text{ g yr}^{-1}$ (ref. 7). The rate of recycling should have been higher in the past due to increased tectonic activity on a younger, hotter Earth, but probably did not exceed the present rate by more than a factor of several. The rate of juvenile degassing of CO_2 could not have been significantly higher than this for more than a few hundred million years, without exceeding the known carbonate inventory in the Earth's crust.

High CO_2 levels could also have occurred if the rate of CO_2 removal by silicate weathering was smaller than it is today. Walker *et al.*¹² have suggested that the atmospheric CO_2 level should rise if the Earth's surface temperature were to fall, due to the temperature dependence of the rainfall rate and the chemical weathering rates of various silicate minerals. According to this model, decreased solar luminosity during the Earth's early history would have resulted in lower average surface temperatures, and hence, higher CO_2 levels (up to $\approx 100 \text{ PAL}$). The increased greenhouse effect due to CO_2 would, in turn, have tended to moderate the temperature drop due to the lower solar luminosity. CO_2 levels could have been further enhanced if outgassing rates were higher than today or if less continental area were exposed to weathering, as suggested by Hargraves¹³.

We have not attempted to develop a self-consistent evolutionary model of CO_2 partial pressure, solar luminosity, and surface temperature. Our purpose here is to calculate the radiative heating due to increased CO_2 levels and compare our results with those of previous studies. As the most detailed treatment has been that of Owen *et al.*⁸, we have adopted the same time history for the changes in solar luminosity and CO_2 partial pressure although this does not imply that we endorse the particular evolutionary sequence proposed by Hart⁹.

The calculation by Owen *et al.*⁸ used a non-grey model and included weak bands of CO_2 which become important for large CO_2 mass paths such as may have existed on the early Earth. A constant relative humidity distribution and a single effective cloud at 6.25 km were specified, consistent with present conditions. Yet one might well expect that the appreciable surface temperature changes brought about by large amounts of CO_2 could alter the distribution of water vapour as well as the cloud height and thickness, all of which affect surface temperature.

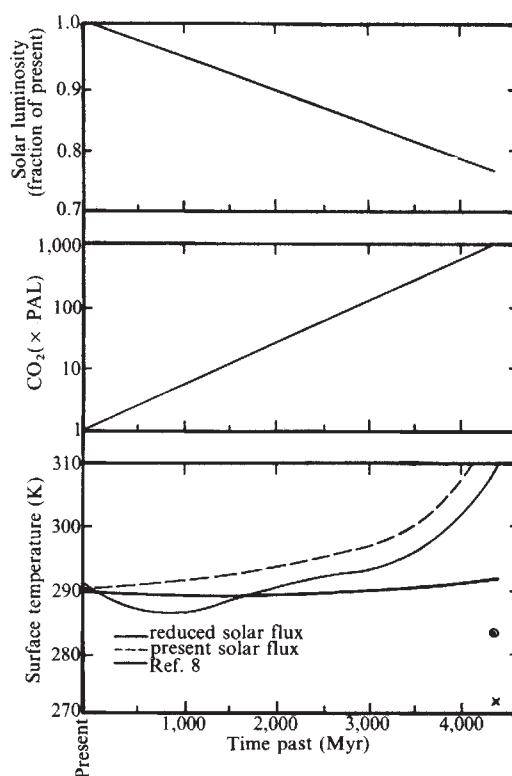


Fig. 1 Mean surface temperature of the primitive Earth for different atmospheric CO_2 concentrations. The temporal variations in solar luminosity and carbon dioxide are from ref. 9. $\odot$ and $\times$ refer to reduced solar flux with CO_2 concentrations of 100 and 1 PAL respectively.

Here we determine the surface temperature for amounts of CO_2 that may have existed in the prebiological atmosphere by including the corresponding changes in water vapour amount and cloud height and thickness.

The model used has been described previously¹⁴ and will only be outlined here. A radiative-convective model is coupled to the water vapour transport. As the temperature profile changes in response to radiative heating or cooling, and latent heating, the distribution of water vapour changes in response to changes in the large scale vertical velocity field and eddy diffusion. Clouds are formed whenever the atmosphere is saturated and the associated latent heat is released to the atmosphere. The number of cloud layers, their altitudes and thicknesses are computed. The total cloud cover is assumed to be 50% for a random cloud distribution.

Figure 1 shows the resultant surface temperature. In addition to showing the results for reduced solar flux, we include calculations with solar flux fixed at the present level as well as the results of Owen *et al.*⁸. The present calculations indicate that, for Hart's CO_2 evolutionary sequence, surface temperature would have varied little over geological time (due to decreased solar flux compensating for larger CO_2 amounts) in contrast to the similar study of Owen *et al.* which indicates a temperature decrease of some 20 K over a 4,250-Myr period. The difference between the two calculations is due primarily to the processes associated with water, that is, the amount of water vapour in the atmosphere, the location of clouds, and the rate at which temperature changes with height. We used a moist adiabatic lapse rate rather than 6.5 K km^{-1} as used by Owen *et al.* which, as shown by Hummel and Kuhn¹⁴ produces a smaller change in surface temperature than does a 6.5 K km^{-1} lapse rate. The cloud top in the present model is calculated to be near 500 mbar while Owen *et al.* fix the cloud top at 6.25 km. This causes their cloud top pressure to be less than in the present study, varying from $\sim 453 \text{ mbar}$ at a 290 K surface temperature to 478 mbar at 310 K, and substantially decreasing the cloud contribution

to the outward radiation, which necessitates a higher surface temperature. Finally, we calculate an increase in atmospheric water vapour of ~30% for 1,000 PAL CO₂ while Owen *et al.*, who assumed constant relative humidity, used four times the present-day water vapour amount. Water vapour does contribute significantly to the greenhouse effect; for example Hansen *et al.*¹⁵ have shown that for a doubling of CO₂, water vapour contributes about as much to the temperature increase as does CO₂.

We conclude that there is still much uncertainty as to the 'greenhouse effect' that could have been produced from increased CO₂ aside from the uncertainty in the CO₂ concentrations. Much of the difficulty concerns water both in the vapour and condensed (cloud) phase. Although our model does represent an improvement in the treatment of atmospheric water vapour, there are still major difficulties. We do fix the surface relative humidity and although the water vapour is calculated when clouds are present, a constant relative humidity assumption is used for clear sky. Yet, one might question whether a constant relative humidity assumption would be valid for atmospheric conditions much different from today, which certainly would have been the case if surface temperatures were even 5K higher than at present. Finally, it is unknown whether the global cloud cover would remain at 50% for different climatological conditions.

While a 'greenhouse effect' is a very plausible mechanism for maintaining the temperature of the early Earth above freezing, the concentrations of CO₂ and water vapour remain very much an uncertainty and the early surface temperature is therefore still unknown. Indeed there is some question concerning the oxidation state of the early atmosphere. Arculus and Delano¹⁶ suggest that the mantle has become progressively more oxidized throughout the Earth's history and that the gases evolved from the original mantle could have been much more reduced than present day volcanic emissions. If CH₄ and CO were present in significant concentrations, they might also have contributed to the greenhouse effect.

This research was supported by NASA grant NSG 7308.

Received 21 June; accepted 27 October 1982.

1. Newman, M. J. & Rood, R. T. *Science* **198**, 1035-1037 (1977).
2. Lowe, D. R. A. *Rev. Earth planet. Sci.* **8**, 145-167 (1980).
3. Sagan, C. & Mullen, G. *Science* **177**, 52-56 (1972).
4. Kuhn, W. R. & Atreya, S. K. *Icarus* **37**, 207-213 (1979).
5. Kasting, J. F. *J. geophys. Res.* **87**, 3091-3098 (1982).
6. Holland, H. D. *Petrologic Studies: A Volume to Honor A. E. Buddington* 447-477 (Geological Society of America, New York, 1962).
7. Walker, J. C. G. *Evolution of the Atmosphere* (Macmillan, New York, 1977).
8. Owen, T., Cess, R. D. & Ramanathan, V. *Nature* **277**, 640-642 (1979).
9. Hart, M. H. *Icarus* **33**, 23-39 (1978).
10. Holland, H. D. *The Chemistry of the Atmosphere and Oceans* (Interscience, New York, 1978).
11. Li, Y. *Am. J. Sci.* **272**, 119-137 (1972).
12. Walker, J. C. G., Hays, P. B. & Kasting, J. F. *J. geophys. Res.* **86**, 9776-9782 (1981).
13. Hargraves, R. B. *Science* **193**, 363-371 (1976).
14. Hummel, J. & Kuhn, W. R. *Tellus* **33**, 372-381 (1981).
15. Hansen, *et al.* *Science* **213**, 957-966 (1981).
16. Arculus, R. J. & Delano, J. W. *Nature* **288**, 72-74 (1980).

Selenium pollution of lakes near the smelters at Sudbury, Ontario

Jerome O. Nriagu & Henry K. Wong

National Water Research Institute, Box 5050, Burlington, Ontario, Canada L7R 4A6

Mining and smelting of Cu-Ni ores at Sudbury, Ontario, release about 2 tonnes of Se to the environment per day. We show here that within a 30-km radius of Sudbury, the Se concentrations in the lake waters range from 0.1 to 0.4 µg l⁻¹; the Se contents of the suspended particulates in the lakes vary from ~2 to 6 µg per g. The Se profiles in the lake sediments generally parallel the history of Cu-Ni production in the district. The present-day rates of Se accumulation in the sediments are found to be 0.3-12 mg m⁻² yr⁻¹. These current deposition rates exceed those of precolonial times by factors of 3-18, and are among the highest recorded anywhere in North America.

Table 1 Se contents of Cu-Ni ores around Sudbury

Mine	Sample no.	Level (ft)	Ore type*	Se concentration (µg per g)
Murray	1	1,650	MS	64
	2	1,650	GPIS	82
	3	1,650	MS	32
Creighton	Cr-1	4,200	IS	26
	Cr-2	3,200	RDS	20
	Cr-3	5,000	GPIS	42
	Cr-4	3,600	MS	65
Garson	Ga-10	2,000	CS	22
Levack	Lev-12	3,000	ISSN	25
Frood	Fr-5	1,000	DSQD	28
	Fr-8	1,000	IS	36
	Fr-9	2,200	RDS	20

Detailed descriptions of the ores, sample locations, relative ore production rates, and so on are given in ref. 11.

* MS, massive sulphide; GPIS, gabbro periodotite inclusion sulphide; IS, interstitial sulphide; RDS, ragged disseminated sulphide; CS, con-torted schist inclusion sulphide in sub-layer norite; DSQD, disseminated sulphide in quartz norite. Details of these ore types are given in ref. 11.

Selenium is an unusual element in terms of its biological function. At low levels, it is an essential nutrient but at slightly higher concentrations it becomes quite toxic^{1,2}. Furthermore, Se typically has an antagonistic effect on the heavy metal (particularly mercury) toxicity, but in certain circumstances can mobilize heavy metals from some of the organs^{3,4}. Little is known about the distribution and possible role of Se in lakes which are highly contaminated with heavy metals.

Selenium replaces sulphur diadochically in sulphides with the result that commercial production of Se is a by-product of the copper (and to a much lesser extent, lead, zinc and iron) industry⁵. The Se contents of representative Cu-Ni ores at Sudbury range from 20 to 80 µg per g (Table 1). During the pyrometallurgical processing, some of the Se is converted to the gaseous phases, but being less volatile and less readily oxidized than the sulphur, it tends to accumulate in the matte, slags and especially the precipitator dusts of the roasting and smelting plants^{6,7}. The awkward chemistry and relatively high volatility of Se and its compounds usually ensure a low overall recovery rate of the Se in the ores of just 5-10% (refs 6, 7).

In 1977, about 17 × 10⁶ tonnes of the Cu-Ni ores were produced in Sudbury⁵. Using the average Se concentration of 40 µg per g, the quantity of Se produced with the ores is estimated to be 680,000 kg. During the same year, the Inco Metals Company plant at Copper Cliff recovered 53,000 kg of Se from the tankhouse slimes⁵. This implies that roughly 630,000 kg (equivalent to 1.7 tonnes day⁻¹), or over 90% of the Se in the ores produced, was lost to the ambient environment. Of these 630 tonnes or so of Se wasted annually in the environment, about 50 tonnes is dispersed through the atmosphere, probably in the form of selenium dioxide⁸. The plumes from the smelter stack at Sudbury (sampled using airborne instruments in both 1979 and 1980 at distances of 1-3 km from the source) indeed show high Se concentrations of 0.1-6.0 µg m⁻³ (W. Chan, personal communication). The rest of the pollutant selenium (580 tonnes) has to be associated with mine tailings, wastewaters and scoria and would constitute a very local source of lake pollution exemplified, for instance, in Kelly Lake (see below).

The lakes sampled are located within a radius of 30 km from the 381-m super stack at Copper Cliff, near Sudbury, Ontario. The water samples were obtained by means of a peristaltic pump with all-plastic tubing. The suspended particulates were recovered from ~1,200 l of water using a Westfalia four-bowl continuous flow centrifuge. At the flow rate of 6 l min⁻¹ used during the study, the centrifuge has a recovery efficiency of 90-95% for particles having diameters >0.25 µm (ref. 9). The particulate samples were kept frozen until freeze-dried.

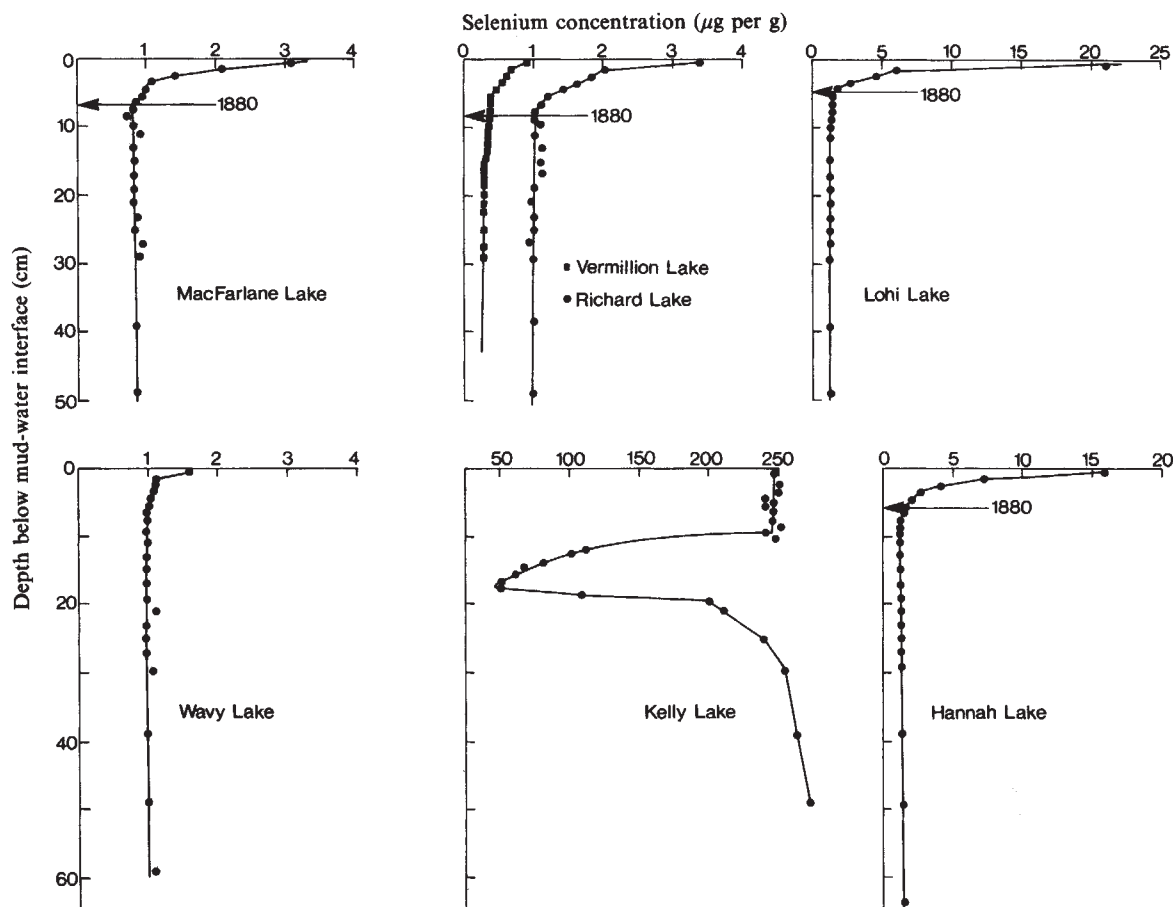


Fig. 1 Representative Se concentrations in sediments of lakes near Sudbury, Ontario.

The lake sediment cores (8-cm diameter) were obtained using a lightweight coring device¹⁰. On retrieval, the cores were quickly sub-sectioned, frozen and subsequently freeze-dried. The ore samples analysed for Se were obtained from Schwarcz who has already characterized them¹¹. The core layers were dated by the ²¹⁰Pb technique using the procedure described previously^{12,13}.

The Se in the sediment, particulate or ore sample was extracted by repeated digestion with nitric, perchloric and hydrofluoric acids in the presence of permanganate and persulphate ions to maintain oxidizing conditions. The Se in the leachate was reduced to the hydride and determined by quartz-tube atomic absorption spectrometry. Details of the method used to determine the total Se content of the samples are given in ref. 14. The error range in the Se measurements is believed to be $\pm 10\%$ of the values reported.

The smelters constitute an obvious point source which may be expected to influence the flux of Se into the lakes in the Sudbury area. The total Se concentrations in representative lakes near Sudbury are given in Table 2. The typical Se concentrations of $0.2\text{--}0.4\ \mu\text{g l}^{-1}$ are higher than the $0.04\text{--}0.21\ \mu\text{g l}^{-1}$ found in some Swedish and Finnish lakes¹⁵. The levels reported in Japanese rivers are considerably lower, being only $0.03\text{--}0.09\ \mu\text{g l}^{-1}$ (ref. 16). The average Se concentrations in the rivers has been estimated to be $0.2\ \mu\text{g l}^{-1}$ (ref. 17), a value that probably needs to be revised downwards.

The accumulation of Se in the suspended particulates in the lakes investigated seems moderate; the levels found range from ~ 3 to $6\ \mu\text{g per g}$ (Table 2). By comparison, the suspended particulates in Rankin and Horseshoe Lakes which are located ~ 240 km south of Sudbury show Se concentrations of only $1\text{--}3\ \mu\text{g per g}$ (Table 2). Few reliable data on the accumulation of Se by the lake sestons are available. The Se contents of planktonic material from the Adriatic Sea average $3.6\ \mu\text{g per g}$ (ref. 18).

The Se profiles in the different lake sediments are summarized in Fig. 1. The Se contents of the precolonial (pre-1880) layers average $\sim 1.0\ \mu\text{g per g}$ dry wt. The precolonial concentrations are thus fairly typical of values reported for other freshwater sediments¹⁹⁻²³. A prominent feature of the Se profiles in Fig. 1 is the pronounced enrichment of this element in the most recent (surficial) sediments. The close similarity with the trace metal profiles²⁴ suggests that the Se enrichment documents recent changes in the flux of Se into the lakes. In other words, the Se profiles are mainly source controlled and not diagenetically determined. The historical changes in the Cu and Ni accumulation in the lake sediments generally parallel the history of Ni and Cu production in the district²⁴. The same historical relationship also seems to hold for Se. Note, for example, that the accelerated flux of Se to the sediments postdate 1890 (see Fig. 1) when the exploitation of ores at Sudbury began²⁴.

The ratios of the Se concentrations in the top 0-1 cm layer to the average concentration in the precolonial layers (defined here as an enrichment factor, EF) fall into two groups (Table 3). The first group represented by Hannah, Lohi and Nelson Lakes show enrichment factors of 13-18 (Table 3). These three lakes had previously been limed to increase their pH (ref. 25), and the lime treatment has apparently orchestrated the removal of Se from the water column. This claim is evident from the EF for the second group of lakes (including Richard, Vermillion and MacFarlane) which range only from 2.8 to 3.5 (Table 3). Most softwater lakes in the area may be expected to have enrichment factors in the 2-5 range, typical of the second group of lakes. Note that the removal of Se by lime treatment suggests that the dominant forms of Se in the lake water are the selenite and selenate; these ions are readily precipitated as poorly soluble salts in alkaline conditions⁷.

The rates of sediment accumulation have been estimated to be 73, 425, 160, 53, 37, 50 and $68\ \text{g m}^{-2}\ \text{yr}^{-1}$ for Richard, Vermillion, MacFarlane, Lohi, Nelson, Kelly and Hannah

Table 2 Se concentrations in the lake waters

Lake	No. of samples	Suspended particulate concentration (mg l ⁻¹)	Total Se concentration (μg l ⁻¹)	Se content of particulate (μg per g dry wt)
Ramsay	6	0.56	0.4	3.6
MacFarlane	4	0.68	0.2	2.6
Vermillion	3	0.61	0.1	3.2
Wavy	2	0.25	0.2	4.8
Nelson	2	0.16	0.2	5.8
Windy	2	0.22	0.1	5.3
Rankin*	4	0.27	0.1	2.7
Horseshoe*	4	0.30	0.1	2.1
Bass*	2	0.81	0.1	1.4

* Data shown for comparison only. Rankin and Horseshoe Lakes are located ~240 km south of Sudbury while Bass Lake is near Orillia, 170 km south-east of Sudbury.

Table 3 Accumulation rates and enrichment factors for Se in lake sediments

Lake	Enrichment factor*	Se flux rate into sediments (mg m ⁻² yr ⁻¹)
Richard	3.4	0.25
Vermillion	2.8	0.40
MacFarlane	3.4	0.49
Lohi	18	1.1
Nelson	13	0.96
Kelly	—	12.4
Hannah	13	1.1

* Enrichment factor defined as ratio of Se concentration in the top 0–1-cm layer to the average concentration in precolonial (deep) layers.

Lakes, respectively²⁴. With these values and the Se concentrations in the top 0–1 cm layers, the rates of Se flux to the lake sediments have been derived (Table 3). The calculated Se deposition rates range from 0.2 to 0.5 mg m⁻² yr⁻¹ for the 'untreated' lakes and average ~1.0 mg m⁻² yr⁻¹ for the lakes that have been limed. We are not aware of any previously reported data on the rate of Se sedimentation in softwater lakes.

The Se concentration and profile in sediments of Kelly Lake are notably different from those of other lakes in the area. The Se concentration drops to a minimum at a depth of ~17 cm below the sediment–water interface (Fig. 1). On the basis of the ²¹⁰Pb profile presented elsewhere²⁴, this feature has been attributed to a past incidence of material dumping into the lake. The Se contents of Kelly Lake sediments are also unusually high, averaging ~250 μg per g in the top 10-cm layers (see Fig. 1). The Se enrichment can be associated with the large fraction of slags and solid wastes in these sediments. Kelly Lake serves as a settling pond for industrial and municipal wastewaters from which it derives substantial quantities of Se-enriched solid wastes. Erosion of the slag heaps which line the lake further enhance the Se concentration in the sediments. The rate of Se accumulation in Kelly Lake sediments of 12 mg m⁻² yr⁻¹ (Table 3), must be among the highest anywhere in North America.

Many studies have shown that Se counteracts the bioaccumulation and toxic effects of Hg (in particular), Ag, Cu, Pb, Cd, Tl, and so on (see ref. 2). It is therefore no coincidence that the mercury levels in fish and zooplankton in many lakes around Sudbury are remarkably low (N. Conroy, personal communication). The biopurification of the organisms with regards to mercury, and presumably other toxic metals, can probably be attributed to the elevated Se concentrations in the lakes. The Se may thus be exercising a greater influence on the biological effects of the high metal contaminants in these lakes than has previously been realized.

We thank B. Van Sickle and N. Arafat for analytical support, and R. D. Coker for assistance with the field work and sample preparation.

Received 5 August; accepted 25 October 1982.

- Rosenfeld, I. & Beath, O. *Selenium: Geobotany, Biochemistry, Toxicity and Nutrition* (Academic, London, 1964).
- Vokal-Borek, H. *Selenium* (Usip Rep. 79–16, University of Stockholm, 1979).
- Magos, L. & Webb, M. in *Biogeochemistry of Mercury in the Environment* (ed. Nriagu, J. O.) 581–599 (Elsevier, Amsterdam, 1979).
- Rudd, J. W. M., Turner, M. A., Townsend, B. E., Swick, A. & Furutani, A. *Can. J. Fish. aquat. Sci.* **37**, 848–857 (1980).
- Canadian Minerals Yearbook 1978* (Energy, Mines and Resources Canada, Ottawa).
- Jennings, P. H. & Yannopoulos, J. C. in *Selenium* (eds Zingaro, R. A. & Cooper, W. A.) 31–85 (Van Nostrand Reinhold, New York, 1974).
- Kudryavtsev, A. A. *The Chemistry and Technology of Selenium and Tellurium* (transl. Elkin, E. M.) (Collet's, London, 1974).
- National Inventory of Sources and Emissions of Selenium* (Rep. EPS 3-AP-77-8, Air Pollution Control Directorate, Ottawa, 1973).
- Williams, J. D. H., Shear, H. & Thomas, R. L. *Limnol. Oceanogr.* **25**, 1–11 (1980).
- Williams, J. D. H. & Pashley, A. E. *J. Fish. Res. Board Can.* **36**, 241–246 (1979).
- Schwartz, H. P. *Can. J. Earth Sci.* **10**, 1444–1449 (1973).
- Nriagu, J. O., Kemp, A. L. W., Wong, H. K. T. & Harper, N. *Geochim. cosmochim. Acta* **43**, 247–258 (1979).
- Nriagu, J. O., Wong, H. K. T. & Coker, R. D. *Water Res.* **15**, 91–96 (1981).
- Agemian, H. & Bedek, E. *Analyt. chim. Acta* **119**, 323–330 (1980).
- Remberger, M. "Förekost och omsättning av selen i inlandsvatten" (Rep. no. B564, Institutet för Vatten och Luftvårdsforskning, 1980).
- Suzuki, Y. & Sugimura, Y. *Jap. J. Limnol.* **42**, 89–91 (1981).
- Kharkar, D. P., Turekian, K. K. & Bertine, K. K. *Geochim. cosmochim. Acta* **42**, 285 (1968).
- Kosta, L. et al. *J. Radioanal. Chem.* **44**, 317–332 (1978).
- Csillag, E. G. thesis, Univ. Toronto (1973).
- Wiersma, J. H. & Lee, G. F. *Envir. Sci. Technol.* **5**, 1203–1206 (1971).
- Speyer, M. R. *Bull. Envir. Contam. Tox.* **24**, 427–432 (1980).
- Heit, M., Klusek, C. S. & Volchok, H. L. *Envir. Int.* **4**, 229–237 (1980).
- Allan, R. J. & Timperley, M. H. in *Prospecting in Areas of Glaciated Terrain*, 87–111 (Institute of Mining and Metallurgy, London, 1975).
- Nriagu, J. O., Wong, H. K. T. & Coker, R. D. *Envir. Sci. Technol.* **16**, 551–560 (1982).
- Yan, N. D. *Water, Air Soil Pollut.* **11**, 43–55 (1979).

Interactive effects of metals and humus on marine phytoplankton carbon uptake

P. B. Ortner, C. Kreader & G. R. Harvey

National Oceanic and Atmospheric Administration, Atlantic Oceanographic and Meteorological Laboratories, 4301 Rickenbacker Causeway, Miami, Florida 33149, USA

Although certain trace metals are essential micronutrients required for growth^{1,2}, elevated concentrations of some of the same metals exert deleterious effects on marine phytoplankton populations^{3–5}. Laboratory studies have indicated that metal toxicity depends on metal ion concentrations rather than total dissolved metals^{6–10}. Although it is believed that bioavailability is largely controlled by the degree to which dissolved trace metals are organically chelated¹¹, it has not been definitively established which organic compounds chelate trace metals in natural seawater¹². In an effort to define ecologically significant interactions between dissolved trace metals and naturally occurring organic matter¹³ we selected marine humus (humic and fulvic acids) as being likely to interact with trace metals in seawater. These compounds, derived from plant and animal sources, are being widely studied for their role in the transport and toxicity of metal ions in terrestrial, aquatic, and marine ecosystems¹⁴ and are known to comprise up to half of the total dissolved organic matter in seawater¹⁵. Guided by a recent hypothesis explaining the structure of marine humus and its geochemical diagenesis¹⁵ a laboratory synthesis of marine fulvic acid was accomplished¹⁶. The resulting material was physically, spectroscopically and chemically identical to one or more natural marine fulvic acids isolated from the Gulf of Mexico. We confirm here that isolated natural marine fulvics and marine fulvics synthesized in the laboratory affect the bioavailability of trace metals to marine phytoplankton.

Short-term net carbon uptake by phytoplankton was determined by a modification of the ¹⁴C-labelled bicarbonate method described in refs 17 and 18. Ultraclean sampling and experimental conditions were maintained essentially as discussed in ref. 19. Immediately after collection, replicate 500-ml sub-aliquots of surface seawater samples were transferred to 500-ml polycarbonate bottles and experimental treatments were inoculated with trace metals or trace metals plus humus. All were allowed to equilibrate for 1 h in either a surface seawater-cooled

Table 1 The effect of added humus on trace metal perturbation of short-term carbon uptake by Gulf Stream phytoplankton

Treatment	% No-addition control ($\pm$ s.e.)	<i>N</i> > Controls*	<i>N</i> < Controls*
Natural fulvic ($300 \mu\text{g l}^{-1}$)	98 ± 5	—	—
Synthetic fulvic ($300 \mu\text{g l}^{-1}$)	110 ± 12	—	—
Copper A (1.8×10^{-8} M)	66 ± 4	—	—
Copper A + natural fulvic	98 ± 8	4	0
Copper A + synthetic fulvic	123 ± 32	4	0
Copper B (1.8×10^{-7} M)	28 ± 6	—	—
Copper B + natural fulvic	39 ± 11	4	0
Copper B + synthetic fulvic	31 ± 12	2	2
Zinc A (5.9×10^{-7} M)	82 ± 4 †	—	—
Zinc A + natural fulvic	95 ± 14 †	1	1
Zinc A + synthetic fulvic	77 ± 10 †	1.5	0.5
Zinc B (5.9×10^{-6} M)	22 ± 4 †	—	—
Zinc B + natural fulvic	30 ± 12 †	1.5	0.5
Zinc B + synthetic fulvic	24 ± 2 †	1	1
Manganese (5.9×10^{-8} M) + copper A (1.8×10^{-8} M)	90 ± 3	—	—
Manganese + copper A + natural fulvic	87 ± 17	2	1
Manganese + copper A + synthetic fulvic	88 ± 17	1.5	2.5

* Values within the range of the controls are scored as ties. Controls are metal-only values.

† The range of replicate values.

deck incubator or a laboratory culture box. Although dynamic equilibrium may not have been attained in 1 h, we feel that experiments of much longer duration would introduce additional physiological artefacts. Samples were then inoculated with ^{14}C -bicarbonate and returned to their incubator for 4–5 h. Light intensities and temperatures were similar in both incubators and were approximately matched to ambient environmental conditions. Following incubation, samples were filtered through Gelman A/E glass fibre filters and particulate ^{14}C measured by scintillation counting. Before counting, samples were fumed for 5 min over concentrated HCl to remove residual inorganic ^{14}C -bicarbonate. All data are corrected to disintegrations per minute (d.p.m.) by the sample-channels-ratio method. In our procedures dark controls never exceed 5% of light values, and time-zero blanks never exceed 1% of light values. Both were neglected in the computations that follow.

Metal stock solutions were prepared from crystalline salts such that with additions of only $50 \mu\text{l}$, final concentrations of from 1.8×10^{-10} M to 5.9×10^{-6} M could be obtained. The natural marine fulvic used was extracted¹⁶ from seawater collected off Galveston, Texas, by a gas lift pumping system²⁰. Artificial fulvic acid was synthesized by allowing ($\pm$)glycerol-1,3-oleate-2-palmitate, an unsaturated triglyceride to auto-oxidize in seawater for 48 h (ref. 16). At least 24 h before initiating an experiment both materials were redissolved in distilled H_2O and stock solutions prepared such that additions of 100–500 μl yielded environmentally reasonable²¹ final concentrations of 300 and $600 \mu\text{g l}^{-1}$.

Two series of experiments were conducted, one in Biscayne Bay (Miami, Florida) and the other within the Straits of Florida in the Gulf Stream east of Miami. Biscayne Bay is a shallow (<5 m depth) marine embayment and is relatively eutrophic. Chlorophyll concentrations in excess of $1.0 \mu\text{g l}^{-1}$ are common. DOC is $2\text{--}3 \text{ mg l}^{-1}$. Near surface total dissolved trace metal concentrations are generally from 1 to $2 \mu\text{g}$ per kg for copper and $0.5\text{--}1.0 \mu\text{g}$ per kg for zinc. In contrast, the Gulf Stream is a highly oligotrophic environment with immeasurably low surface macro nutrients (nitrate, phosphate, silicate and ammonia). Surface chlorophyll values rarely, if ever, exceed $0.1 \mu\text{g l}^{-1}$ at the sampled latitude. DOC is $<1.0 \text{ mg l}^{-1}$ while near surface total dissolved trace metal concentrations are generally from 10 to 20 ng per kg for zinc and were measured on this occasion to be 21 ng per kg for copper. To avoid introducing artefactual variability, Biscayne Bay sampling was always done at low tide. Biscayne Bay experiments were run in the laboratory. To eliminate prejudicial delay between sample collection and experimental incubation, Gulf Stream experiments were run at sea aboard the RV *Virginia Key*.

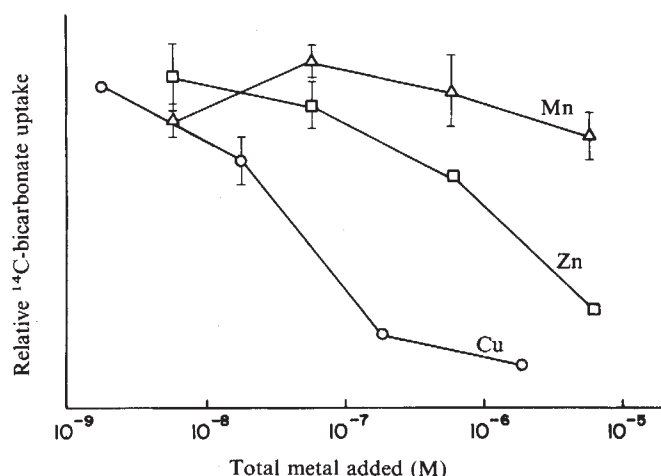


Fig. 1 Response of Gulf Stream phytoplankton to added trace metals. Data are expressed as percentages of the short-term carbon uptake of replicate no addition controls. Error bars equal the range of replicate determinations. Metal augmentations resulting from each addition are expressed in molar units.

Additions of fulvic acids alone ($600 \mu\text{g l}^{-1}$), whether natural or synthetic, were apparently benign and did not affect Biscayne Bay carbon uptake (data not given). Based on preliminary experiments two copper concentrations (1.8×10^{-7} and 1.8×10^{-6} M) and one zinc concentration (5.9×10^{-6} M) were chosen as inhibitory. The effect of these additions was determined with and without co-addition of fulvic acids. Although the selected metal concentrations were inhibitory, at neither copper concentration nor at the single zinc concentration did fulvic acids significantly reduce depression of carbon uptake.

The response of Gulf Stream phytoplankton to low level additions differed for the three metals tested (Fig. 1). All metals were tested on sub-aliquots of the same water sample. Copper was inhibitory at tested concentrations well below those for zinc. For example at the two highest copper additions tested, carbon uptakes were 20 and 11% of control values, while at the two highest zinc additions uptakes were 68 and 27% of control values. Manganese was never inhibitory. Gulf Stream phytoplankton carbon uptake was significantly depressed at zinc and copper concentrations an order of magnitude lower than had been the case in Biscayne Bay.

Once again, fulvic acids alone, regardless of origin, did not affect carbon uptake (Table 1). Based on Fig. 1, two copper concentrations (1.8×10^{-8} and 1.8×10^{-7} M) and two zinc concentrations (5.9×10^{-7} and 5.9×10^{-6} M) were chosen as inhibi-

tory. As no manganese concentration tested produced any effect, a concentration within the range of selected copper values was chosen to investigate Mn-Cu interactions. The effect of these additions was then determined with and without addition of fulvic acids. The selected metal concentrations were inhibitory and significantly depressed carbon uptake to a similar degree as in the preliminary experiments depicted in Fig. 1. In contrast to Biscayne Bay, co-addition of fulvics and metals produced an appreciable effect. At the lower copper concentration tested the addition of fulvic acid, either natural or synthetic, alleviated carbon uptake inhibition. At the higher copper concentration tested these differences were less appreciable. Although the confidence intervals about the means for both fulvic types overlap that of the copper-only treatment, it is suggestive that both means exceeded the copper-only mean and that six of eight individual determinations exceeded all four copper-only values. Neither fulvic affected zinc-induced inhibition. The addition of manganese also alleviated copper toxicity at the lower copper concentration. Confidence intervals about the means of copper-only and copper-manganese treatments do not overlap. The addition of fulvics to the copper-manganese treatment did not restore uptake to control levels.

Copper was found to be considerably more pernicious than zinc. Gulf Stream phytoplankton populations were adversely affected at considerably lower concentrations of both metals than were their Biscayne Bay counterparts. These results are consistent with previous studies^{12,22}. We cannot determine with certainty whether this regional contrast was due to a fundamental difference in Biscayne Bay and Gulf Stream water chemistry or to a difference in the inherent metal sensitivity of geographically separate phytoplankton populations. Recent results suggest both factors can be important but that explanations may differ for zinc and copper²³⁻²⁷. It has been hypothesized that one mechanism of copper toxicity may be competitive inhibition of manganese uptake²⁸. This would explain why manganese largely ameliorated copper inhibition in our Gulf Stream samples. Were a large excess of manganese present, reducing copper free-ion concentration by the addition of a chelator would not be expected to have a significant effect.

Natural marine humus has only recently been isolated in quantities sufficient to permit experimentation^{15,20}. It has been demonstrated with anodic stripping voltammetry (ASV)²⁹ that marine humus interacts with both copper and zinc at natural levels of those elements and at the pH of seawater³⁰. It has also been demonstrated that marine humus can reduce manganese (IV) oxide in a light-dependent reaction and may, therefore, be critical in maintaining dissolved manganese in a reduced form within the euphotic zone³¹. These experiments, conducted in conditions similar to our own, indicate that isolated humic matter redissolved in seawater appears to retain its original metal-binding characteristics. ASV and copper bioassays, however, can yield very different results^{32,33}. Purely chemical evidence cannot resolve whether or not marine humus controls trace metals speciation in the environment. Previous biological experiments similar to ours have used synthetic chelators such as EDTA⁶⁻¹⁰, terrestrial humus³⁴, or uncharacterized algal exudates in conditioned growth media^{35,36}. The data presented earlier represent the first biological indications that isolated natural marine humus affects trace metal bioavailability. It behaves exactly as one would expect of a chelator in that it significantly decreases free ion concentration in proportion to both the amount of trace metals added and ambient environmental chelation capacity³⁷. Our data also strongly support a recently proposed mechanism for the structure and diagenesis of marine humus¹⁵. They establish that synthetic humus prepared from the postulated precursor compounds is not only a spectroscopic analogue of natural marine humus, but is also a functional analogue.

We thank D. Boran for providing the synthetic and natural humic substances used in this study and the crew and captain of the RV *Virginia Key* for assistance at sea. This research was funded by NOAA's Office of Marine Pollution Assessment.

Received 23 August; accepted 5 November 1982.

1. Provasoli, L., McLaughlin, J. J. A. & Droop, M. R. *Arch. Mikrobiol.* **25**, 392-426 (1957).
2. Guillard, R. L. & Ryther, J. H. *Can. J. Microbiol.* **8**, 229-239 (1962).
3. Thomas, W. H., Seibert, D. L. R. & Takahashi, M. *Mar. Sci. Commun.* **3**, 331-354 (1979).
4. Thomas, W. H., Hollebaugh, J. T., Seibert, D. L. R. & Wallace, G. T. *Jr Mar. Ecol. Prog. Ser.* **2**, 213-220 (1980).
5. Davies, A. G. & Sleep, J. A. *Nature* **277**, 292-293 (1979).
6. Sunda, W. G. & Guillard, R. L. *J. mar. Res.* **34**, 511-529 (1976).
7. Anderson, D. M. & Morel, F. M. M. *Limnol. Oceanogr.* **23**, 283-295 (1978).
8. Jackson, G. A. & Morgan, J. J. *Limnol. Oceanogr.* **23**, 268-282 (1978).
9. Morel, N. M. L., Reuter, J. G. & Morel, F. M. M. *J. Phys.* **14**, 43-48 (1978).
10. Gavis, J., Guillard, R. L. & Woodward, B. L. *J. mar. Res.* **39**, 315-333 (1981).
11. Sunda, W. G. & Ferguson, R. L. in *Proc. NATO/ARI Symp. on Trace Metals in Seawater* (ed. Wong, C. S.) (Plenum, New York, in the press).
12. Huntsman, S. A. & Sunda, W. G. in *The Physiological Ecology of Phytoplankton* (ed. Morris, I.), 285-328 (University of California Press, 1980).
13. Atwood, D. K. *EOS* **63**, 70 (1982).
14. Saar, R. A. & Weber, J. H. *Envir. Sci. Technol.* **16**, 510A-517A (1982).
15. Harvey, G. R., Boran, D. A., Chesal, L. A. & Tokar, J. M. *Mar. Chem.* **12**, 1-14 (1983).
16. Harvey, G. R. & Boran, D. A. *EOS* **63**, 62 (1982).
17. Steeman-Nielsen, E. *J. Cons. Int. Explor. Mar.* **18**, 117-140 (1952).
18. Strickland, J. D. H. *Fish. Res. Board. Canada Bull No.* 167 (1972).
19. Fitzwater, S. E., Knauer, G. A. & Martin, J. A. *Limnol. Oceanogr.* **27**, 544-551 (1982).
20. Tokar, J. M., Harvey, G. R. & Chesal, L. A. *Deep-Sea Res.* **28**, 1395-1399 (1981).
21. Huizenga, D. L. & Kester, D. R. *Limnol. Oceanogr.* **24**, 145-157 (1979).
22. Wood, A. M. in *Proc. 21st Hanford Life Sciences Symp. Biological Availability of Trace Metals* (Elsevier, Amsterdam, in the press).
23. Brand, L. E., Sunda, W. G. & Guillard, R. L. *EOS* **63**, 112 (1982).
24. Sunda, W. G. & Huntsman, S. *EOS* **63**, 113 (1982).
25. Ortner, P. B., Piotrowicz, S. R., Chesal, L., Berberian, G. & Ferguson, R. L. *EOS* **63**, 88 (1982).
26. Fisher, N. S. & Frood, D. *Mar. Biol.* **59**, 85-93 (1980).
27. Fisher, N. S., Jones, G. J. & Nelson, D. M. *Mar. Chem.* (in the press).
28. Sunda, W. G., Barber, R. T. & Huntsman, S. A. *J. mar. Res.* **39**, 567-586 (1981).
29. Piotrowicz, S. R., Springer-Young, M., Puig, J. A. & Spencer, M. J. *Analyt. Chem.* **54**, 1367-1371 (1982).
30. Piotrowicz, S. R., Harvey, G. R., Springer-Young, M., Courant, R. A. & Boran, D. A. in *Proc. NATO/ARI Symp. on Trace Metals in Seawater* (ed. Wong, C. S.) (Plenum, New York, in the press).
31. Sunda, W. G., Huntsman, S. A. & Harvey, G. R. *Nature* (in the press).
32. Gachter, R., Lum-Shue-Chan, K. & Chan, Y. K. *Schweiz. Z. Hydrol.* **35**, 252-261 (1973).
33. Srna, R. F., Garrett, K. S., Miller, S. M. & Thum, A. B. *Envir. Sci. Technol.* **14**, 1482-1486 (1980).
34. Prakash, A. *Fertility of the Sea* (ed. Costlow, J. D. Jr) 351-358 (Gordon & Breach, London, 1971).
35. Fisher, N. S. & Fabris, J. G. *Mar. Chem.* **11**, 245-255 (1982).
36. Gnassia-Barelli, M., Romeo, M., Laumond, F. & Pesando, D. *Mar. Biol.* **47**, 15-19 (1978).
37. Spencer, M. J. & Carpenter, J. *EOS* **63**, 112 (1982).

Partial bivoltinism may cause alternating sex-ratio biases that favour eusociality

Jon Seger*

Population Biology Group, School of Biological Sciences, University of Sussex, Falmer, Brighton BN1 9QG, UK

Fisher¹ showed that frequency-dependent selection will stabilize balanced sex ratios of parental investment², under a wide range of mating systems and patterns of mortality. But many actual ratios of investment are biased. Some of these can be explained by models³⁻¹⁴ in which the life history lacks certain kinds of symmetry that are present in Fisher's model. For example, where brothers compete directly for matings, female biases are expected and found^{3,9}. Here I describe a model for the evolution of sex ratios in species having two generations per year with an asymmetrical pattern of overlap between the generations. Many insects have partially bivoltine life histories of this kind. Two predictions emerge from an analysis of the model. First, one of the two generations produced each season will be female-biased, and the other will be male-biased. Second, eusociality will arise more often in haplodiploid species that overwinter as inseminated females than it will in those that overwinter as larvae or as unmated adults. Data on the sex ratios of bivoltine solitary Hymenoptera and on the phylogenetic distribution of eusociality are consistent with these predictions. Thus the model shows one particular way in which ecological and genetic variables may interact to encourage the evolution of eusociality in some species and to discourage it in others.

* Present address: Museum of Zoology, University of Michigan, Ann Arbor, Michigan 48109, USA

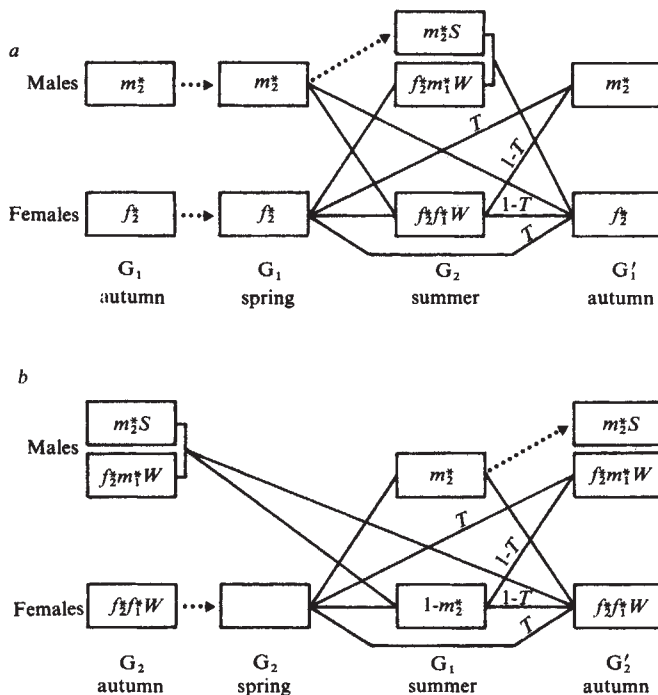


Fig. 1 Partially bivoltine life histories. Each block represents a cohort of males or females at a given season. Solid lines represent paths of genetic transmission between parents and offspring, for a haplodiploid genetic system. Dotted arrows represent direct transfers of individuals from one season to another. Terms inside the blocks give the relative sizes of the cohorts present at any given time, as explained in the text. (Note that $f_1^* = 1 - m_1^*$.) The spring cohorts in *a* have the same relative sizes as do those of the previous autumn, despite the fact that they may be absolutely smaller owing to mortality. The spring cohort of females in *b* is left blank because there are no males present. The factors T and $1 - T$ indicate the relative contributions made to the overwintering generation by females belonging to the two previous generations. Male contributions to female offspring are identical to those of their mates. *a*, 'Sphecoid' model. When $S > 0$, males belonging to G_1 contribute to the next two generations, no matter what the value of T . *b*, 'Halictine' model. When $S > 0$, males belonging to G_1 contribute to the next two generations if $T = 0$, and to the next three generations if $T > 0$. This appears to be the main reason why, at a given value of S/W , sex ratios become more biased as T increases under the halictine model, but less biased under the sphecoid model (Fig. 2).

A partially bivoltine life history is shown in Fig. 1*a*. Males and females overwinter as last-instar larvae. They complete development, emerge, and mate in late spring or early summer. Females of this first or 'spring' generation (G_1) then produce offspring that develop directly, emerging after only a few weeks to form the second or 'summer' generation (G_2). Offspring produced by G_2 enter diapause as last-instar larvae and spend the winter in that state, emerging the following year as members of G_2' . In a purely bivoltine life history these generations would be completely discrete, in the sense that all members of one generation would be offspring of members of the previous generation. But two kinds of overlap occur in the life history being considered here. First, some of the G_1 males survive to take part (along with G_2 males) in the mating of G_2 females. Second, G_1 females produce some offspring who, instead of developing directly and emerging in G_2 , enter diapause and emerge the following year in G_2' . Thus males and females of the first generation (G_1) may have offspring in the next two generations (G_2 and G_2'). Females belonging to G_1 mate with one or more G_1 males, shortly after emerging in the spring, but they do not mate with G_2 males.

Three critical features of this life history will be quantified by the parameters S , T , and W . S is the probability that an adult G_1 male survives to take part in the G_2 mating. T is the

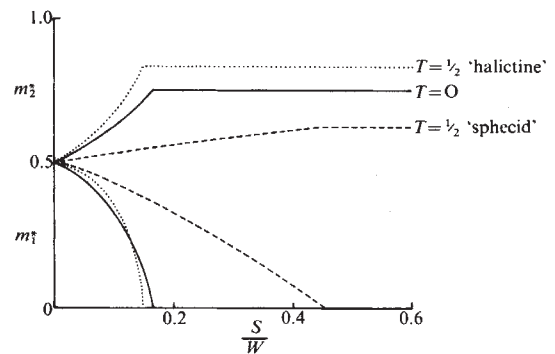


Fig. 2 Evolutionarily stable ratios of investment (m_1^* and m_2^*) as functions of S/W , for haplodiploid genetic systems. Solid lines ($T = 0$) represent cases in which there is no overlap of generations on the female side. Broken lines ($T = \frac{1}{2}$, 'sphecoid') represent cases in which half of the next overwintering generation is produced by females of the current overwintering generation, under the 'sphecoid' model (Fig. 1*a*), and dotted lines ($T = \frac{1}{2}$, 'halictine') represent these cases under the 'halictine' model (Fig. 1*b*). Although m_1^* and m_2^* depend on the parameters S , T and W , it is most convenient to express the general stability conditions as functions of m_2^* and T . Under the sphecoid model,

$$m_1^* = \frac{3 - 4m_2^* - T}{2(2 - 2m_2^* - T)}, \quad \frac{S}{W} = \frac{(1 - m_2^*)(2m_2^* - 1)(1 + T)}{2m_2^*(2 - 2m_2^* - T)}$$

up to the point at which m_1^* falls to zero. Under the halictine model,

$$m_1^* = \frac{1}{2} - \frac{(2m_2^* - 1)(1 - T)}{4(1 - m_2^*)}, \quad \frac{S}{W} = \frac{(2m_2^* - 1)(1 - T^2)}{4m_2^*}$$

Results for diploid genetic systems are identical to these, except that the expressions for S/W are missing a factor $(1 + T)/2$ on their right-hand sides. Thus if $T = 0$, S/W must be twice as large under diploidy as under haplodiploidy to stabilize a given pair of investment ratios (m_1^* and m_2^*).

proportion of G_2' produced by females belonging to G_1 . W is the average number of surviving G_2 offspring produced by each G_1 female. Note that S and T determine the amount of overlap, or departure from pure bivoltinism, on the male and female sides respectively. When $T = 1$ the model becomes univoltine, and S is meaningless. It is assumed that males and females of any given generation suffer equal egg-to-adult mortality rates, and that females mate randomly with the males alive at the time of their emergence.

Given S , T and W , it is possible to find the evolutionarily stable sex 'ratios' of investment m_1^* (the proportion of male investment in G_2 , which is that produced by G_1 females giving rise to G_2 offspring) and m_2^* (the proportion of male investment in G_2' , which is that produced by G_2 females and by G_1 females giving rise to G_2' offspring). This was done by constructing diploid and haplodiploid genetic models with two loci. One locus has alleles inducing m_1^* and its alternative m_1 , and the other has alleles inducing m_2^* and m_2 . A female's sex ratio phenotype is determined by her genotype at one locus or the other, as appropriate. Expressions for m_1^* and m_2^* were derived from conditions for local stability¹³⁻¹⁵ of the recurrence equations for the gene frequencies at each locus in each sex, near fixation for the alleles inducing m_1^* and m_2^* . Details of the analysis will be published elsewhere. Results were checked numerically, by iterating the full system of recurrence equations for the genotype frequencies, and were found in all cases to hold for intermediate gene frequencies as well as for the boundary at fixation. These results are illustrated for the haplodiploid case in Fig. 2. Values of m_1^* and m_2^* are plotted as functions of the ratio S/W for $T = 0$ and $T = \frac{1}{2}$. At any given value of T , G_1 becomes increasingly male-biased ($m_2^* > \frac{1}{2}$) and G_2 becomes increasingly female-biased ($m_1^* < \frac{1}{2}$) as S/W increases, up to the point at which m_1^* falls to zero. The results for diploidy are identical to those for haplodiploidy, except that different combinations of S/W and T are required to produce a given

pair of investment ratios, as explained in the legend to Fig. 2. Werren and Charnov⁸ analysed a diploid model with a structure very similar to that of this model when $T = 0$ (no overlap on the female side). They found a similar dependence of the sex ratio biases on the survival of G_1 males, and the same limits of m_1^* and m_2^* (0 and $\frac{1}{2}$, respectively).

One of the unusual features of this model is that the ratios of investment depend in a non-fisherian way on rates of survival and fecundity. S is a rate of survival after the period of parental investment, and W is the product of a fecundity and an egg-to-adult survival. There is no sense in which the average ratio of investment over both generations equals $\frac{1}{2}$, except in the trivial case in which both m_1^* and m_2^* equal $\frac{1}{2}$. The easiest way to see this is to note that the model does not specify the fecundity of G_2 females. Thus in principle, the overall ratio of investment could take almost any value between m_1^* and m_2^* .

There are sex ratio data for many bivoltine solitary Hymenoptera, but I have been able to find only 10 species for which the sex ratios of first and second generations are reported separately (Table 1). These sex ratios are not quantitative estimates of the ratios of investment, because the average weights of males differ from those of females⁵. More importantly, there is no way to predict exactly what the ratios of investment ought to be in these species, because there are no demographic data with which to estimate the parameters S , T and W , and with which to rule out or estimate the effects of other possible causes of bias, such as local mate competition³. But it is reasonable to suppose that S commonly exceeds zero. Individual male aculeates have often been seen 3–5 weeks after marking^{16–18}, and Evans¹⁹ followed some males for as long as 7 weeks during an intensive study of one population of the digger wasp *Philanthus gibbosus*. A first-generation male in a typical bivoltine species would need to survive only 4–8 weeks (as an adult) to mate with females of both the first and second generations, and doing so could greatly increase his fitness. In any species where S does exceed zero, the sex ratio of G_1 should exceed that of G_2 . Nine of the 10 species in Table 1 show the expected difference ($P < 0.01$, sign test). The two generations differ significantly from each other in five species.

All of the species in Table 1 have life histories of the form discussed above and illustrated in Fig. 1a. This life history will be referred to as the 'sphecid' model because it is almost universal among bivoltine members of the wasp family Sphecidae. But a different bivoltine life history occurs in some aculeate taxa, most notably in the bee subfamily Halictinae. In the 'halictine' model (Fig. 1b), inseminated females hibernate through the winter, emerging in the spring to begin production of a summer brood. This first direct-developing generation (G_1) corresponds to the G_1 of the sphecid model. Its members mate and produce a second direct-developing generation (G_2), which emerges in late summer or early autumn. After mating, G_2 females go into hibernation. No males survive the winter. The halictine and sphecid models are formally identical as long as they remain purely bivoltine on the female side ($T = 0$). But different generations overwinter in the two models (G_1 in the sphecid model, and G_2 in the halictine model). Thus they differ when $T > 0$, because in both models it is the overwintering females who contribute to the two subsequent generations. When $T > 0$, ratios of investment can be even more biased under the halictine model than they are under the sphecid model, for given values of S and W (Fig. 2).

Under either model, a summer female might emerge to find her mother still alive and reproducing. When might such a daughter be selected not to reproduce on her own, but instead to work for her mother, as a eusocial 'helper at the nest'? Assume that the daughter can help her mother produce an additional B offspring (who will be the daughter's siblings), but that in doing so she forfeits C offspring of her own. Let $K = B/C$. If S were equal to zero, both m_1^* and m_2^* would be equal to $\frac{1}{2}$, and it would follow from a well known inclusive fitness^{20,21} argument that helping would be favoured only if K were greater than one. In fact, $K > 1$ is the condition for positive selection

of helping at any sex ratio, under haplodiploidy or diploidy, if the same sex ratio is produced by all females in all generations²². Hamilton²¹ and Trivers and Hare⁵ showed that helping can evolve where K is less than one, in a haplodiploid species, if females with helpers produce an excess of female offspring (sisters of the helpers, related to them by $\frac{1}{2}$) and a deficit of male offspring (brothers of the helpers, related to them by $\frac{1}{4}$), while females without helpers produce an excess of males (sons, to which they are related by $\frac{1}{2}$). But under the halictine model described here, helping can be favoured at $K < 1$ even if females with helpers produce exactly the same sex ratio as do females without helpers, within any given generation. Assume that $S > 0$ and $T > 0$. Summer females belong to G_1 and thus make female-biased ratios of investment. If the mother of a summer female is alive and reproducing, she too is making a female-biased ratio of investment, because at that point in the season she too is contributing to the overwintering generation (G_2). Under these circumstances the critical value of K is less than one (potentially as low as $2/3$) for the initial spread of helping in a haplodiploid species, but not in a diploid species. In the sphecid model a male-biased overwintering generation is produced, and the critical value of K is greater than one, for a haplodiploid species. An explicit derivation of this result will be published elsewhere. Sex ratio biases would evolve in multi-voltine versions of the bivoltine models discussed here, but the biases would be smaller. Other things being equal, then, eusociality should arise most easily in partially bivoltine haplodiploid species in which the overwintering generation consists of mated females.

Table 1 Sex ratios of solitary bivoltine wasps and bees

Species	M_1	M_2	P	Ref.
Pompilidae				
<i>Dipogon sayi</i>	0.36 (33)	0.55 (11)	(0.3)	27
Eumenidae				
<i>Ancistrocerus antilope</i>	0.33 (12)	0.84 (31)	0.002	27
<i>A. antilope</i>	0.09 (76)	0.67 (103)	0.0005	29
<i>A. catskill</i>				
<i>albophaleratus</i>	0.30 (254)	0.85 (40)	0.0005	27
<i>Euodynerus foraminatus</i>	0.50 (133)	0.52 (56)	(0.3)	29
<i>E. foraminatus</i>	0.47 (552)	0.52 (318)	(0.2)	J.W.S.
<i>E. leucomelas</i>	0.32 (44)	0.54 (28)	(0.06)	27
Sphecidae				
<i>Pemphredon lethifer</i>	0.32 (111)	0.59 (533)	0.0005	28
<i>Passaloecus ithacae</i>	0.29 (28)	0.31 (70)	(0.8)	27
Megachilidae				
<i>Megachile relativa</i>	0.32 (115)	0.54 (331)	0.001	29
<i>M. mendica</i>	0.70 (33)	0.54 (76)	(0.2)	29
<i>M. inermis</i>	0.17 (6)	0.88 (17)	0.01	29

All of the species in this table normally make nests in hollow twigs or other crevices and all have 'sphecid' life histories (Fig. 1a). Most of the data are from artificial trap-nests offered throughout one or more complete nesting seasons. It is known that the inside diameter of a trap-nest can affect the sex ratio of the offspring produced in it (ref. 12). In most of the studies a range of diameters was offered, and the same distribution of vacant nests was maintained throughout the season. Thus even if the sex ratios were affected by the sizes of the offered nests, there is no reason to suppose that the sign of the difference between the two generations was affected. M_1 and M_2 are the observed proportions of males, corresponding to m_1^* and m_2^* . If a study continued for more than 1 yr, data from all years have been combined. Numbers following the sex ratios are total sample sizes. P is the significance level of the difference between M_1 and M_2 (χ^2 test, 1 d.f.). Two independent sets of data are shown for two of the eumenid species. Both sets of data for *Ancistrocerus antilope* are from populations in Ontario, Canada, but these populations are from different localities and are separated in time by 16 yr. The first set of data for *Euodynerus foraminatus* is from a population in Ontario, and the second is from a population in Massachusetts (J. W. Stubblefield, personal communication). All six eumenid populations in the table show differences of the expected kind between the sex ratios M_1 and M_2 . But it is interesting that the three populations of *Ancistrocerus* (representing two species) show very large differences, while the three populations of *Euodynerus* (also representing two species) show only modest differences. Three populations are referred to here by names different from those used in the original publications: *Ancistrocerus antilope* (= *A. catskill catskill* in ref. 27), *Euodynerus leucomelas* (= *Rygiellium leucomelas* in ref. 27), and *Pemphredon lethifer* (= *Cemonus lethifer* in ref. 28).

There is a higher rate of bivoltinism among insect species of the temperate regions of the world than there is among those of the tropics, where reproduction often continues through most of the year. Thus contact between mothers and daughters should be relatively common in tropical species, but sex ratio biases should be relatively common (and relatively large) in temperate species. This suggests that in diploid groups, where sex ratio biases offer no advantage to potential helpers, eusociality may be more likely to evolve in tropical than in temperate species. The only diploid eusocial insects are the termites, almost all of whom are tropical²³. But in haplodiploid groups, where sex ratio biases may offer a substantial advantage to potential helpers, eusociality should be comparatively likely to evolve in temperate species. There have been at least 11 origins of eusociality within the aculeate Hymenoptera and most of the eusocial groups have temperate and tropical members²³. Outside the tropics, most of the primitively eusocial species overwinter as inseminated females²³⁻²⁵, and many of these species belong to groups whose solitary members do so as well. In contrast, most of the temperate solitary aculeates overwinter as larvae or as unmated adults of both sexes. These taxonomic associations will be reviewed elsewhere.

If seasonal variation of the sex ratio can favour eusociality, seasonal and nonseasonal populations belonging to a geographically widespread species might exhibit different levels of sociality. Michener and Bennett²⁶ describe what may be such a situation in the bee *Halictus ligatus*. In Trinidad the species is active throughout the year. New nests may be established by solitary females at any time. Thus alternating sex ratio biases are not expected, because the population is not synchronously bivoltine. Daughters frequently remain in their mother's nest, but they appear to act more as independent reproductives than as workers. Most of them eventually mate and develop enlarged ovaries, and all members of the colony forage. In New York and Ontario the species overwinters as inseminated females. New nests are established in the spring, and there are two or three generations per year. Thus alternating sex-ratio biases are expected, and other things being equal, helping behaviour should be favoured more often than it is in the tropics. In fact, most nests become fully eusocial. Foragers seldom mate or show signs of ovarian development, and their mother does not forage. It would be interesting to know whether this pattern occurs in any other species (or series of closely related species), and if so, whether their sex ratios vary in the expected way.

I thank J. W. Stubblefield for unpublished data on *Euodynerus foraminatus*, D. B. Brown, M. G. Bulmer, B. Charlesworth, D. Charlesworth, T. H. Clutton-Brock, M. C. Day, G. R. Else, H. E. Evans, P. H. Harvey, R. W. Longair, J. R. Malcolm, J. Maynard Smith, J. Pickering, W. M. Post, O. W. Richards, J. W. Stubblefield, J. H. Werren and N. Yamamura for helpful advice and criticism. This work was supported by a NATO Postdoctoral Fellowship from the US National Science Foundation.

Received 21 May; accepted 11 October 1982.

1. Fisher, R. A. *The Genetical Theory of Natural Selection* (Oxford University Press, 1930).
2. Trivers, R. L. in *Sexual Selection and the Descent of Man 1871-1971* (ed. Campbell, B.) (Aldine, Chicago, 1972).
3. Hamilton, W. D. *Science* **156**, 477-488 (1967).
4. Trivers, R. L. & Willard, D. E. *Science* **179**, 90-92 (1973).
5. Trivers, R. L. & Hare, H. *Science* **191**, 249-263 (1976).
6. Charnov, E. L. & Bull, J. *Nature* **266**, 828-830 (1977).
7. Clark, A. B. *Science* **201**, 163-165 (1978).
8. Werren, J. H. & Charnov, E. L. *Nature* **272**, 349-350 (1978).
9. Werren, J. H. *Science* **208**, 1157-1159 (1980).
10. Bulmer, M. G. & Taylor, P. D. *Nature* **284**, 448-449 (1980).
11. Bull, J. J. *Hereditas* **46**, 9-26 (1981).
12. Charnov, E. L., Los-den Hartogh, R. L., Jones, W. T. & Assem, J. van den *Nature* **289**, 27-33 (1981).
13. Charnov, E. L. *Am. Nat.* **112**, 317-326 (1978).
14. Charnov, E. L. *Am. Nat.* **113**, 465-480 (1979).
15. Travis, C. C., Post, W. M., DeAngelis, D. L. & Perkowski, J. *Theor. pop. Biol.* **18**, 16-30 (1980).
16. Simon Thomas, R. T. & Poorter, E. P. R. *Tijdschr. Ent.* **115**, 141-151 (1972).
17. Alcock, J. et al. *Zool. J. Linn. Soc.* **64**, 293-326 (1978).
18. Gwynne, D. T. *Behav. Ecol. Sociobiol.* **7**, 213-225 (1980).
19. Evans, H. E. *Anim. Behav.* **21**, 302-308 (1973).
20. Hamilton, W. D. *J. theor. Biol.* **7**, 1-16 (1964).

21. Hamilton, W. D. *Ann. Rev. Ecol. Syst.* **3**, 193-232 (1972).
22. Craig, R. J. *J. theor. Biol.* **87**, 55-70 (1980).
23. Wilson, E. O. *The Insect Societies* (Harvard University Press, 1971).
24. Michener, C. D. *The Social Behavior of the Bees* (Harvard University Press, 1974).
25. Stephen, W. P., Bohart, G. E. & Torchio, P. F. *Biology and External Morphology of Bees* (Agricultural Experiment Station, Oregon State University, Corvallis, 1969).
26. Michener, C. D. & Bennett, F. D. *Univ. Kansas Sci. Bull.* **51**, 233-260 (1977).
27. Fye, R. E. *Can. Ent.* **97**, 716-744 (1965).
28. Danks, H. V. *Trans. R. ent. Soc. Lond.* **122**, 323-399 (1970).
29. Longair, R. W. *Evolution* **35**, 597-600 (1981).

β -Endorphin regulates lordosis in female rats by modulating LH-RH release

D. J. S. Sirinathsinghji, P. E. Whittington, A. Audsley & H. M. Fraser*

Neuroendocrine Research Unit, Animal Physiology Division, Agricultural Research Council, Meat Research Institute, Langford, Bristol BS18 7DY, UK

* Medical Research Council, Reproductive Biology Unit, Centre for Reproductive Biology, Edinburgh EH3 9EW, UK

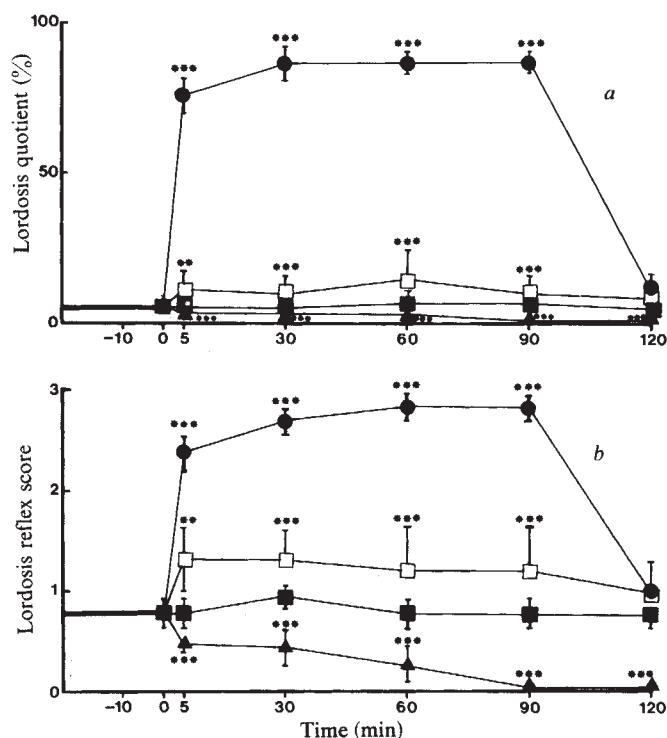
Several lines of evidence have implicated the endogenous opioid peptides in the regulation of masculine sexual behaviour¹⁻¹⁰. However, although the opioid related peptides α -melanocyte stimulating hormone (α -MSH)^{11,12} and adrenocorticotropin (ACTH)¹³ have been shown to affect lordosis behaviour in the female rat, there is as yet no evidence for a role of the endogenous opiates in the regulation of female sexual behaviour. We present here evidence that the endogenous opiates in the mesencephalic central grey (MCG) are involved in the control of lordosis behaviour in the female rat.

The MCG is rich both in β -endorphin-like^{14,15} and enkephalin-like¹⁶⁻¹⁸ immunoreactivity and opiate receptors^{16,19} and contains heavy descending hypothalamic projections²⁰ including luteinizing hormone-releasing hormone (LH-RH) axons^{21,22}. Both electrical stimulation²³ and LH-RH infusion^{24,25} into the MCG facilitate lordosis behaviour in oestrogen-primed female rats. LH-RH has also been shown to have neurotropic effects when microiontophoretically applied to the MCG²⁶. As the suppressive effects of endogenous opiates on luteinizing hormone (LH) secretion may be due to suppression of LH-RH from hypothalamic sites^{27,28}, we considered it possible that endogenous opiates could also regulate lordosis behaviour by modulating the release of LH-RH.

The infusion of naloxone (200 ng) into the MCG induced a prompt facilitation of the lordosis response (Fig. 1). Facilitation occurred within 5 min and lasted for about 90 min. Higher concentrations of naloxone up to 500 ng had no further facilitatory effects and lower doses (<200 ng) were not as effective (data not shown). An equivalent volume of saline infused into the MCG had no effect and an equivalent amount of naloxone infused into the superior colliculus had no effect on lordosis behaviour. Two strategies were used to determine if this facilitation of lordosis was due to naloxone-induced LH-RH release. First, it was envisaged that LH-RH released from LH-RH terminals in the MCG could be neutralized with anti-LH-RH antiserum with a resulting diminution or abolition of the lordosis response, and second, the released LH-RH would have no facilitatory effects on lordosis if LH-RH receptors on those MCG neurones sensitive to LH-RH and involved in the control of lordosis are blocked by a specific LH-RH antagonist analogue. The results show that pretreatment of the MCG site with anti-LH-RH antiserum diminished the naloxone-induced potentiation of lordosis throughout the period studied but did not completely abolish lordosis (Fig. 1). Pretreatment with the LH-RH antagonist analogue not only inhibited the naloxone-induced lordosis but completely abolished lordosis behaviour within 90 min (Fig. 1). The effect was still seen 12 h later.

Fig. 1 The effect of saline (■—■), naloxone (●—●), naloxone + anti-LH-RH antiserum (□—□) and naloxone + LH-RH antagonist [D-Phe², Pro³, D-Phe⁶]LH-RH (▲—▲) infusions into the mesencephalic central grey on lordosis behaviour in oestrogen-primed rats assessed in terms of lordosis quotient (a) and lordosis reflex score (b). Each time point denotes the mean $\pm$ s.e.m. of 12 rats. Asterisks (** $P < 0.01$, *** $P < 0.001$, Student's *t*-test) denote statistically significant differences between groups (naloxone versus saline; naloxone versus naloxone + anti-LH-RH antiserum; naloxone versus naloxone + LH-RH antagonist analogue).

Methods: Stainless steel 29-gauge guide cannulae were chronically implanted bilaterally into the MCG in ovariectomized Sprague-Dawley female rats (190–220 g) using the stereotaxic coordinates³⁸ AP: -5.4; L: ± 0.5 ; V: 4.25. Two days later each animal was injected subcutaneously with 5 μ g of oestradiol benzoate in arachis oil and tested 96 h later for sexual receptivity by pairing with a sexually vigorous male. Only those females showing a low to moderate degree of lordosis were used for the experiments. All infusions were done with 31-gauge stainless steel tubes fitted into the guide cannulae and protruding 0.5 mm from the tip of the latter. Solutions of the substances used in the study were prepared in saline in the following concentrations: naloxone hydrochloride (Endo), 400 μ g ml⁻¹; human β -endorphin (Cambridge Research Biochemicals, CRB), 500 μ g ml⁻¹; Met-enkephalin (CRB), 1 mg ml⁻¹; LH-RH (NIH), 200 μ g ml⁻¹; LH-RH antagonist analogue, [D-Phe², Pro³, D-Phe⁶]LH-RH (CRB), 500 μ g ml⁻¹. Samples (1.0 ml) of a high titre anti-LH-RH (no. 94) antiserum³⁹ were lyophilized and then resuspended in 10 μ l saline. The binding capacity of the undiluted antiserum³⁹ was 12 μ g ml⁻¹. Recent studies (unpublished) show that the antibody formation is directed towards the C-terminal end of the LH-RH molecule and was highly specific for LH-RH; it showed no cross-reactivity (<0.01%) with corticotropin releasing factor (CRF), thyrotropin releasing hormone (TRH), somatostatin (SRIF), substance P or β -endorphin. Anti-LH-RH antiserum and LH-RH antagonist analogue were each given 10 min before naloxone. Each solution was prepared immediately before use and was infused in a volume of 0.5 μ l over a period of 1 min using a Harvard infusion pump (model 975). Two sets of inner cannulae and Hamilton microsyringes were used for the simultaneous delivery of each solution into the bilateral guide cannulae. The internal infusion cannulae were left in place for 1 min after the end of the infusion before withdrawal. The female was then paired with a sexually vigorous male and the lordosis reflex score (LS) and the lordosis quotient (LQ) (number of lordoses $\times$ 100/number of mounts) were assessed. The lordosis reflex score was determined using the scale 0 (no vertebral dorsiflexion) to 3 (full dorsiflexion). The LS and LQ were assessed immediately after the end of the infusion and then every 30 min. The LQ was calculated as per cent lordoses per 20 mounts. The average of 20 ratings of the reflex score was collected independently by three investigators. At the end of the experiment, the animals were killed and the brains removed for histological confirmation of the placement of cannula tips.



No inhibitory effect on lordosis was seen when normal sheep antiserum was infused into the MCG nor when anti-LH-RH antiserum or the LH-RH antagonist analogue were infused into the superior colliculus. Although all the animals repeatedly showed potentiation of lordosis behaviour with naloxone, there were no consistent indications of proceptive behaviour. Only 5 out of 12 animals showed hopping and darting movements but none showed ear-wiggling. Similar observations were noted in studies with subcutaneous injections of LH-RH^{29,30}.

The use of naloxone in the above study could not provide precise information as to which of the endogenous ligands of the opiate receptors are actually involved in the modulation of LH-RH release and thus regulation of lordosis behaviour, as it acts equally well on each of the different opiate receptors. We therefore investigated the effects of β -endorphin and Met-enkephalin when infused into the MCG. β -Endorphin (250 ng) infused into the MCG completely abolished lordosis behaviour in the oestrogen-primed female rat within 2 h (Fig. 2), with inhibition lasting 7 h. The suppressive effect of β -endorphin on lordosis behaviour could be overcome not only by pretreatment of the animal with an intraperitoneal injection of naloxone (2.0 mg per kg) but by a relatively small dose of LH-RH (100 ng) infused directly into the MCG (Fig. 2). LH-RH promptly overcame the β -endorphin blockade by potentiating lordosis behaviour to near maximal levels and thus confirmed the effects of the decapeptide²⁵. This response lasted $\sim$ 90 min even in the presence of β -endorphin. In contrast to β -endorphin, Met-enkephalin (500 ng) infused into a different group of animals had no inhibitory effects on lordosis behaviour. Larger doses (10 μ g) also had no effect (Fig. 2). The effects of these two opiate peptides on lordosis were also assessed in the oestrogen-progesterone-primed ovariectomized female rat.

The peptide was infused into the MCG of animals ($n = 10$) injected 44 h previously with 5 μ g of oestradiol benzoate. After the opiate peptide infusion, they were immediately injected subcutaneously with 500 μ g of progesterone and tested for sexual receptivity 4 h later. Compared with the saline control, the infusion of β -endorphin (250 ng) significantly suppressed ($P < 0.005$, Student's *t*-test) but did not abolish the lordosis response whereas Met-enkephalin (10 μ g) had no significant inhibitory effect. The oestrogen-progesterone-treated ovariectomized female rat displays sexual receptivity which is associated with an LH surge and an increase in hypothalamic LH-RH^{31–34}. Such an increase in LH-RH with progesterone may also have taken place in the MCG to overcome any total abolition of the lordosis response with β -endorphin which was observed in the animal primed only with oestrogen (Fig. 2).

The results of these studies taken together give further but more conclusive evidence for the role of LH-RH in the regulation of lordosis behaviour. Of more relevance, however, the studies demonstrate that the endogenous opiates, specifically β -endorphin in the MCG, may control sexual receptivity by regulating the release of LH-RH. Although both Met-enkephalin neurones and terminals are present in the MCG^{16–18}, this opiate peptide may not be involved in the mesencephalic neuronal circuitry for the neuroendocrine regulation of lordosis. Despite evidence that Met-enkephalin can inhibit LH-RH release from mediobasal hypothalamic (MBH) slices^{35,36}, our data support more recent data obtained in endocrine studies with specific antisera to the opiates suggesting that β -endorphin rather than Met-enkephalin is involved in the regulation of LH secretion and thus LH-RH release³⁷.

The mechanisms by which β -endorphin fibres in the MCG¹⁴ interact with LH-RH terminals to regulate the release of LH-

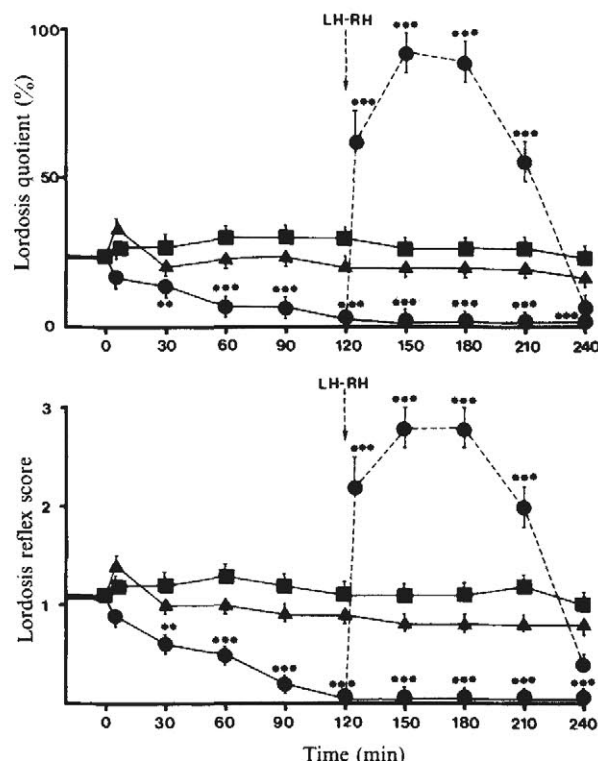


Fig. 2 The effect of saline (■—■), Met-enkephalin (▲—▲), β -endorphin (●—●) and β -endorphin + LH-RH (◆—◆) infusions into the MCG on lordosis behaviour in oestrogen-primed rats. Each time point denotes the mean $\pm$ s.e.m. of 10 rats. Asterisks (** $P < 0.01$, *** $P < 0.001$, Student's t -test) denote statistically significant differences between groups (β -endorphin versus saline; LH-RH versus β -endorphin; LH-RH versus saline).

27. Drouva, S. V., Epelbaum, J., Tapia-Arancibia, L., La Plante, E. & Kordon, C. *Neuroendocrinology* **32**, 163–167 (1981).
28. Kalra, S. P. *Endocrinology* **109**, 1805–1810 (1981).
29. Moss, R. L. & McCann, S. M. *Science* **181**, 177–179 (1973).
30. Pfaff, D. W. *Science* **182**, 177–179 (1973).
31. Johnston, P. G. & Davidson, J. M. *Neuroendocrinology* **28**, 155–159 (1979).
32. Kalra, S. P., Kalra, P. S., Chen, C. L. & Clemens, J. A. *Acta endocr.* **89**, 1–9 (1978).
33. Lu, K.-H. & Yen, S. S. C. *Endocrinology* **106**, 867–870 (1980).
34. Levine, J. E. & Ramirez, V. D. *Endocrinology* **107**, 1782–1790 (1980).
35. Rotsztein, W. H., Drouva, S. V., Patton, E. & Kordon, C. *Nature* **274**, 281–282 (1978).
36. Drouva, S. V., Epelbaum, J., Tapia-Arancibia, L., Laplante, E. & Kordon, C. *Neuroendocrinology* **32**, 163–167 (1981).
37. Schulz, R., Wilhelm, A., Pirke, K. M., Gramsch, C. & Herz, A. *Nature* **294**, 757–759 (1981).
38. Pellegrino, L. J., Pellegrino, A. S. & Cushman, A. J. *A Stereotaxic Atlas of the Rat Brain* (Plenum, New York, 1979).
39. Clayton, R. N., Popkin, R. M. & Fraser, H. M. *Endocrinology* **110**, 1116–1123 (1982).

Anxiogenic and non-anxiogenic benzodiazepine antagonists

Lia Prado de Carvalho, Gisela Grecksch, Georges Chapouthier & Jean Rossier

Laboratoire de Physiologie Nerveuse, Centre National de la Recherche Scientifique, 91190 Gif-sur-Yvette, France

Benzodiazepines are widely used anxiolytic and anticonvulsant drugs, and brain receptors for these drugs have been characterized by Möhler and Okada¹ and Squires and Braestrup². Recently, substances that antagonize benzodiazepine binding to brain receptors have been discovered^{3,4}. These benzodiazepine antagonists were shown to block the central effects of benzodiazepines and particularly their anticonvulsive properties^{3,5,6}. Two such antagonists, Ro 15-1788 (an imidazodiazepine) and methyl β -carboline-3-carboxylate (β -CCM), have recently been shown to have different intrinsic pharmacological properties. β -CCM, injected into baboons, cats, mice and rats, is a convulsant^{6–8}, whereas Ro 15-1788 lacks such an activity. Thus, the separation of convulsant and non-convulsant antagonists has been proposed⁸. We suggest here that a sub-classification of antagonists is also valid at the behavioural level, based on a conflict model in mice. We show that Ro 15-1788 and β -CCM antagonize the anxiolytic effect of benzodiazepines. In addition, we find that, when injected alone, Ro 15-1788 has no anxiogenic effects while β -CCM has anxiogenic properties. We therefore propose that β -CCM is an anxiogenic convulsant benzodiazepine antagonist and that Ro 15-1788 is a non-anxiogenic non-convulsant benzodiazepine antagonist.

Conflict generated by simultaneous punishment and reward of a previously learned task is widely accepted as a good animal model for anxiety^{9–12}. Punishment is usually an electric shock and reward may be a food pellet for food-deprived animals or a water drop for water-deprived animals. Classical conflict models include periods of reward only and periods of conflict (reward plus punishment). The performance of the task is measured separately for each period and is always reduced during the conflict period. In our conflict model, mice were deprived of food until they attained 80% of their free feeding weight. They were then trained in a Skinner box to press a lever for a food pellet. When stable pressing rates of approximately 8 presses per min were obtained, mice were submitted to 15-min daily sessions divided as follows: during the first and last 5-min periods, each lever press resulted in a food pellet (non-conflict); during the central 5-min period, each lever press was similarly rewarded but also concomitantly punished with an electric foot-shock (conflict). After 2 weeks of training with this 15-min paradigm, mice showed stable day to day performances. Figure 1 shows the pressing rates during each 5-min period in five subsequent sessions. Lever-pressing rates during conflict were always reduced to less than 10% of the first 5 min, except for diazepam treatment days. During the last 5 min lever-pressing rates were again high. The duration of each lever

RH and hence lordosis are unclear. Recent data suggest that opiate receptors may be present on LH-RH terminals^{27,28}. Our data indicate that naloxone may overcome a presynaptic β -endorphin inhibition of LH-RH release from terminals in the MCG and that there may be functional axo-axonic interactions between β -endorphin and LH-RH neurones in the vicinity of MCG neurones sensitive to LH-RH for the regulation of lordosis behaviour.

Received 8 September; accepted 8 November 1982.

1. Meyerson, B. J. & Terenius, L. *Eur. J. Pharmac.* **42**, 191–192 (1977).
2. Goldstein, A. & Hansteen, R. W. *Archs gen. Psychiat.* **34**, 1179–1180 (1977).
3. Pelligrini-Quarantotti, B. M., Corda, M. G., Paglietti, E., Biggio, G. & Gessa, G. L. *Life Sci.* **23**, 673–678 (1978).
4. Murphy, M. R., Bowie, D. L. & Pert, C. B. *Soc. Neurosci. Abstr.* **5**, 470 (1979).
5. Myers, B. M. & Baum, M. J. *Pharmac. Biochem. Behav.* **10**, 615–618 (1979).
6. Pelligrini-Quarantotti, B. M. et al. *Experientia* **35**, 524–525 (1979).
7. McIntosh, T. K., Vallano, M. L. & Barfield, R. J. *Pharmac. Biochem. Behav.* **13**, 435–442 (1980).
8. Myers, B. M. & Baum, M. J. *Pharmac. Biochem. Behav.* **12**, 365–370 (1980).
9. Meller, R. E., Keverne, E. B. & Herbert, J. *Pharmac. Biochem. Behav.* **13**, 663–672 (1980).
10. Szechtman, H., Hershkowitz, M. & Simantov, R. *Eur. J. Pharmac.* **70**, 279–285 (1981).
11. Thoday, A. J., Wilson, C. A. & Everard, D. *Physiol. Behav.* **22**, 447–450 (1979a).
12. Thoday, A. J., Wilson, C. A. & Everard, D. *Psychopharmacology* **74**, 153–156 (1981).
13. Wilson, C. A., Thoday, A. J. & Everard, D. *Horm. Behav.* **13**, 293–300 (1979).
14. Watson, S. J., Akil, H., Richard, C. W. & Barchas, J. D. *Nature* **275**, 226–228 (1978).
15. Bloom, F. E., Battenberg, E., Rossier, J., Ling, N. & Guillemin, R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1591–1595 (1978).
16. Hokfelt, T., Elde, R., Johansson, O., Terenius, L. & Stein, L. *Neurosci. Lett.* **5**, 25–31 (1977).
17. Hokfelt, T. et al. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3081–3085 (1977).
18. Sar, M., Stumpf, W. E., Miller, R. J., Chang, K. J. & Cuatrecasas, P. *J. comp. Neurol.* **182**, 17–37 (1977).
19. Atweh, S. F. & Kuhar, M. J. *Brain Res.* **129**, 1–12 (1977).
20. Krieger, M. S., Conrad, L. C. A. & Pfaff, D. W. *J. comp. Neurol.* **83**, 785–816 (1979).
21. Silverman, A. J. & Krey, L. C. *Brain Res.* **157**, 233–246 (1978).
22. Liposits, Zs. & Setaloi, G. *Neurosci. Lett.* **20**, 1–4 (1980).
23. Sakuma, Y. & Pfaff, D. W. *Am. J. Physiol.* **237**, R278–R284 (1979).
24. Riskind, P. & Moss, R. L. *Brain Res. Bull.* **4**, 203–205 (1979).
25. Sakuma, Y. & Pfaff, D. W. *Nature* **283**, 566–567 (1980).
26. Samson, W. K., McCann, S. M., Chud, L., Dudley, C. A. & Moss, R. L. *Neuroendocrinology* **31**, 66–72 (1980).

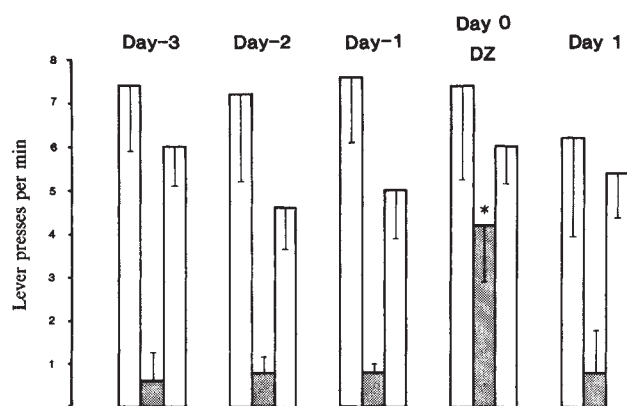


Fig. 1 Mouse conflict model and the effect of diazepam. Lever pressing rates (median $\pm$ interquartile range) of 13 mice on subsequent daily sessions of a conflict test. Each session was subdivided into three periods: a central 5-min period of conflict (shaded bars) which was preceded and followed by 5-min periods of reward only (non-conflict, open bars). The reward was a 20 mg Noyes precision food pellet; punishment was a 1 s 60 μ A electric foot-shock. Days -3, -2 and -1 correspond to the last three sessions before the drug test. Lever pressing rates during the conflict periods were lower than during the non-conflict periods in all sessions. On day 0, diazepam (DZ, 1 mg per kg, s.c.) was administered 10 min before the beginning of the session. * Indicates a significant effect of diazepam on the pressing rates during conflict compared with pressing rates during the conflict period of the previous and following days (Wilcoxon $T = 1$ and $T = 1.5$, respectively; $P = 0.005$). Diazepam had no effect on pressing rates during non-conflict periods.

press was recorded graphically and we observed that during conflict each lever press was much shorter (0.2 s) than during non-conflict periods (6 s).

In this model, diazepam (DZ, 1 mg per kg, subcutaneously, s.c.), the prototype anxiolytic drug, enhanced lever-pressing rates during conflict periods (Fig. 1, day 0, DZ). Diazepam treatment did not change the lever-pressing duration. Twenty-four hours after diazepam injection, pressing rates were similar to pre-drug baseline levels. Using each mouse as its own control, the results of Fig. 1 were transformed as shown in Fig. 2. With this representation it can be seen that on the diazepam day the lever-pressing rate during conflict had a median 520% of the previous day values. This increase was not seen when diazepam administration was followed by subsequent administration of Ro 15-1788 (20 mg per kg, s.c.) or β -CCM (1 mg per kg, s.c.). Thus, following a combined treatment of either diazepam and Ro 15-1788 or diazepam and β -CCM, lever-pressing rates remained low during conflict (in spite of diazepam). Similarly to results with diazepam alone, pressing rates during non-conflict periods were unaltered. These results demonstrate that the two benzodiazepine antagonists block the anxiolytic effect of diazepam.

A possible anxiogenic effect of benzodiazepine antagonists has been suggested by Dorow¹³, File *et al.*¹⁴, Petersen *et al.*¹⁵ and ourselves¹⁶ and this was the underlying basic hypothesis for the second phase of our study. If anxiolytic drugs enhance lever-pressing rates during conflict, anxiogenic drugs could have the opposite effect, that is, to reduce lever-pressing rates even further. To quantify the possible decrease, we had to use mice with high pressing rates during the conflict period. These mice were nevertheless sensitive to the shock. Indeed, the decrease in duration of each lever press during conflict was clear and comparable (0.2 s) with that obtained in mice with low lever-pressing rates. In this model, Ro 15-1788 alone (20 mg per kg, s.c.) had no effect on pressing rates during conflict or non-conflict periods (Fig. 3). This indicates that in this model Ro 15-1788 has no intrinsic activity³.

To study the behavioural effects of β -CCM, we had to choose a dose which did not induce convulsions. We observed that a dose of 1 mg per kg, s.c., does not cause tonic-clonic convul-

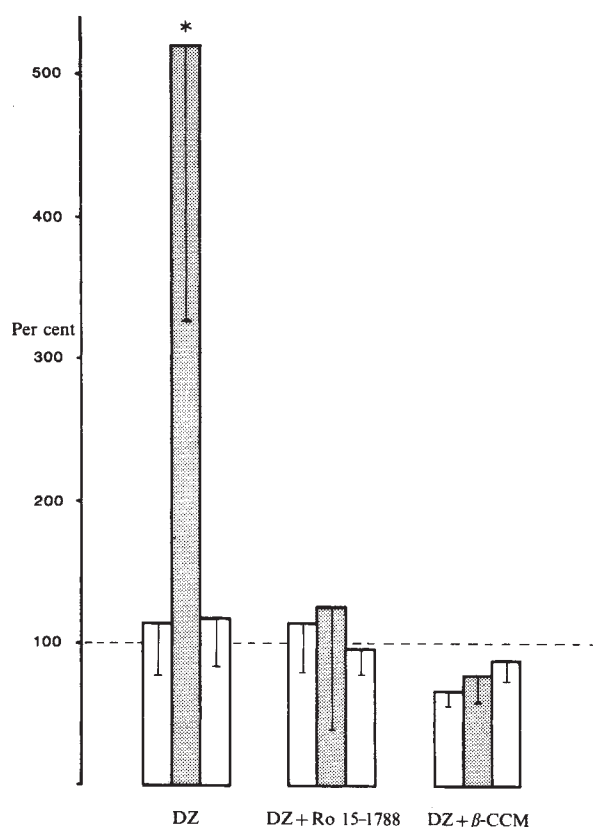


Fig. 2 The effect of diazepam on conflict and its reversal by Ro 15-1788 and β -CCM. Each mouse was used as its own control; therefore, pressing rates during each 5-min period of the session were calculated as a percentage of the corresponding previous day value. Results are expressed as median percentage $\pm$ interquartile. Diazepam results of Fig. 1 were thus transformed in this figure. * Indicates significant effect of diazepam on conflict (Wilcoxon $T = 1$, $T = 1.5$; $P = 0.005$). Diazepam (1 mg per kg, s.c.) was always injected 10 min before the beginning of the session. Ro 15-1788 was suspended in saline with a drop of Tween 80 and administered at a dose of 20 mg per kg, s.c., 2 min before the beginning of the session ($n = 7$). β -CCM was dissolved in 0.01 M HCl (1 mg per 100 μ l) and diluted with saline. A dose of 1 mg per kg was administered 3 min before the beginning of the session ($n = 9$). Volume injected was 0.05 ml per 10 g mouse. Shaded bars, conflict; open bars, non-conflict.

sions. Moreover, in mice implanted with recording electrodes in the cortex and hippocampus, this dose did not induce electrical seizure. A few spikes were occasionally observed in these areas during the 30 min following the injection.

This dose of 1 mg per kg of β -CCM also had no effect on pain threshold, food ingestion or motor coordination. Pain threshold was measured on a 54 $^{\circ}$ C hot plate. β -CCM-treated mice, similarly to saline-treated mice, showed no increase or decrease in latency to lick the paw or to jump. Food-deprived mice injected with β -CCM 2, 5 or 10 min before presentation of their usual food began consumption as did uninjected animals. β -CCM-treated mice also did not differ from controls when submitted to the rotarod test; they could stay on the rotating cylinder (4 r.p.m.) for > 1 min.

Mice trained on the conflict paradigm were then tested with the 1 mg per kg dose of β -CCM. While behaving similarly to the saline-treated animals in the rotarod, food control and pain threshold tests, mice injected with 1 mg per kg β -CCM showed behavioural suppression when placed in the Skinner box. During the first 5 min, these mice remained in a corner of the box and did not press the lever for food. They then progressively resumed pressing and the 15-min daily session could be started. However, compared with previous day pressing rates, a median 37% reduction was observed during the first 5 min of the session. Interestingly, this reduction was significantly stronger

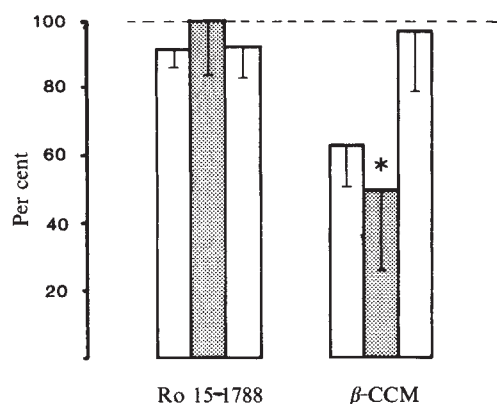


Fig. 3 Anxiogenic effects of β -CCM and non-anxiogenic effect of Ro 15-1788. Animals with a 4 min^{-1} pressing rate were used. Each mouse was used as its own control. Results are expressed as median percentage $\pm$ interquartile of previous day values. The punishment used was a 1 s, $60 \mu\text{A}$ electric foot-shock. Ro 15-1788 and β -CCM solutions and administration were as in Fig. 2. β -CCM caused a significant reduction in pressing rates in both the initial 5 min of non-conflict and the 5 min of conflict (Wilcoxon $T = 1.5$, $T = 4$, respectively, $n = 13$; $P = 0.005$). * Indicates a significantly stronger reduction during conflict than during the first 5 min (Wilcoxon $T = 21$, $n = 13$; $P = 0.05$) or last 5 min ($T = 1$; $P = 0.005$). No significant difference was seen for mice treated with Ro 15-1788 alone ($n = 7$). Shaded bars, conflict; open bars, non-conflict.

during the conflict period (median 50%), the effect postulated for anxiogenic drugs. The initial behavioural inhibition and reduction in lever pressing during the first 5 min of the test could be interpreted as a general anxiogenic effect of β -CCM in periods that precede conflict. This anxiogenic interpretation is supported by the complete recovery of previous day baseline levels during the last 5 min of the session, a period when all risk of shock has passed. This complete recovery could not be explained by a decreased effectiveness of β -CCM because we have observed that β -CCM has a prolonged action and its effect can be demonstrated for more than 30 min. Thus, in a rotarod test, a dose of 3 mg per kg diazepam will impair the ability of mice to stay on a rotating cylinder. Such a deficit could be reversed within 2–30 min of a single injection of β -CCM (1 mg per kg).

We have shown here that two benzodiazepine antagonists block the anticonflict effect of diazepam. One of these, β -CCM, is a convulsant benzodiazepine antagonist but also had pronounced behavioural effects when injected alone and in subconvulsive doses. This compound induced behavioural suppression and reduced lever pressing rates for food during a conflict period. Together, such effects have been interpreted as an index of the anxiogenic properties of β -CCM. In contrast to β -CCM, the other benzodiazepine antagonist tested, Ro 15-1788, was devoid of any activity of its own on the same behavioural paradigm. It seems that β -CCM is an anxiogenic benzodiazepine antagonist whereas Ro 15-1788 is a non-anxiogenic benzodiazepine antagonist.

It is well accepted that benzodiazepine receptors are coupled to the receptor for γ -aminobutyric acid (GABA). Experiments by Study and Barker¹⁷ have indicated that benzodiazepine agonists increase the frequency of Cl^- channel opening generated by GABA binding. Such an increase in GABA efficacy may be the molecular mechanism underlying the anticonvulsant and anxiolytic actions of benzodiazepine agonists. To explain the convulsant and anxiogenic effect of β -CCM, we propose that binding of β -CCM to the benzodiazepine receptor has an opposing action, decreasing the effectiveness of GABA inhibition and thereby generating convulsion and anxiety. It thus seems that β -CCM is pharmacologically not a classical antagonist. In contrast with β -CCM, Ro 15-1788 appears to be a classical antagonist whose binding to the receptor does not

include a conformational change but only displacement of other ligands.

We thank Dr W. Haefely for a gift of Ro 15-1788 and Dr R. Dodd for a gift of β -CCM. L.P.C. was supported by CAPES (Brazil). G.G. was supported by an IBRO-INSMER fellowship. This work was supported by INSERM (PRC medicament 121022), by CNRS (ATP Pharmacologie des récepteurs des neuromédiateurs), by PIRMED and by the Esther A. and Joseph Klingenstein Fund (New York).

Received 27 September; accepted 1 November 1982.

1. Möhler, H. & Okada, T. *Science* **198**, 849–851 (1977).
2. Squires, R. F. & Braestrup, C. *Nature* **266**, 732–734 (1977).
3. Hunkeler, W. *et al. Nature* **290**, 514–516 (1981).
4. Braestrup, C., Nielsen, M. & Olsen, C. E. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2288–2290 (1980).
5. Rommelspacher, H. *et al. Eur. J. Pharmac.* **70**, 409–416 (1981).
6. Braestrup, C., Schmichen, R., Neef, G., Nielsen, M. & Petersen, E. N. *Science* **216**, 1241–1243 (1982).
7. Oakley, N. R. & Jones, N. R. *Neuropharmacology* **21**, 587–589 (1982).
8. Valin, A., Dodd, R. H., Liston, D. R., Potier, P. & Rossier, J. *Eur. J. Pharmac.* **85**, 93–97 (1982).
9. Geller, I., Kulak, J. T. & Seifter, J. *Psychopharmacologia* **25**, 112–116 (1962).
10. Geller, I. *Archs int. Pharmac. Ther.* **149**, 243–247 (1964).
11. Tye, N. C., Iversen, S. D. & Green, A. R. *Neuropharmacology* **18**, 689–695 (1979).
12. Cook, L. & Davidson, A. B. in *The Benzodiazepines* (eds Garattini, S., Mussini, E. & Randall, L. O.) 327–345 (Raven, New York, 1973).
13. Dorow, R. *Abstr. 13th Collegium int. Neuropsychopharmac. Congr.*, Jerusalem, 175 (1982).
14. File, S. E., Lister, R. G. & Nutt, D. J. *Neuropharmacology* **21**, 1033–1037 (1982).
15. Petersen, E. N., Paschelke, G., Kehr, W., Nielsen, M. & Braestrup, C. *Eur. J. Pharmac.* **82**, 217–221 (1982).
16. Prado de Carvalho, L., Grecksch, G., Chapouthier, G. & Rossier, J. *Behav. Brain Res.*, *Abstr. Eur. Brain Behav. Soc. Meet.*, Parma (in the press).
17. Study, R. E. & Barker, J. L. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7180–7184 (1981).

Muscle pioneers: large mesodermal cells that erect a scaffold for developing muscles and motoneurons in grasshopper embryos

Robert K. Ho, Eldon E. Ball & Corey S. Goodman

Department of Biological Sciences, Stanford University, Stanford, California 94305, USA

During embryonic development, muscles differentiate in the appropriate places and motoneurone growth cones find the appropriate muscles; both events occur concurrently and with remarkable specificity. What are the cellular interactions that orchestrate this coordinated development of nerve and muscle? In the development of vertebrate skeletal muscles, motoneurone growth cones arrive in the periphery along stereotyped routes and enter the appropriately located masses of mesodermal cells usually before differentiated muscle fibres appear and before the masses cleave into separate muscles^{1–3}. We find that a similar sequence of events occurs in the grasshopper embryo. We are interested in how mesodermal cells become organized into the appropriate muscles and what guides motoneurone growth cones to their appropriate targets. Fortunately, in the grasshopper embryo the mesodermal cells in the periphery and motoneurons in the central nervous system (CNS) are large, accessible and in many cases individually identifiable from early in their development. We report here the discovery of a class of large mesodermal cells, which we call muscle pioneers, that arise early in development when the embryonic environment is relatively simple and distances short. By their growth and association with particular sites along the ectoderm, the muscle pioneers appear to erect a scaffold for later developing muscles and motoneurone growth cones.

The cell bodies of the muscle pioneers are visible in the living embryo viewed with Nomarski optics, and portions of their morphology can be seen in the fixed embryo viewed with scanning electron microscopy (SEM). Their complete morphology is more readily studied following staining with a monoclonal antibody (I-5 mab) followed by horseradish peroxidase (HRP)-labelled secondary antibody⁴, or by intracellular injections of tracer dyes (HRP and Lucifer yellow, LY) and sub-

sequent immunocytochemistry with an antibody to one of these dyes (anti-LY)⁵. We have used these methods to study the cellular events underlying muscle development and neuromuscular specificity in the body wall and limb buds of the grasshopper embryo. The I-5 mab stains a cytoplasmic antigen in a subset of neurones and mesodermal cells in the grasshopper embryo^{4,6}. The muscle pioneers described here stain with the I-5 mab early in their development.

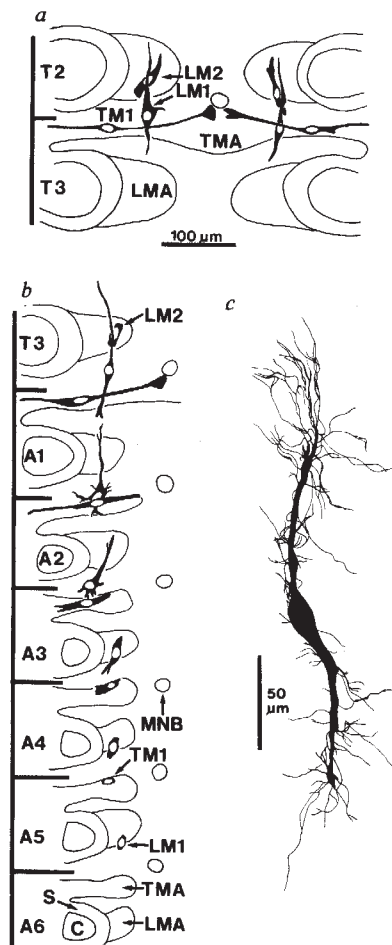


Fig. 1 Muscle pioneers in the body wall of the grasshopper embryo, as revealed by immunocytochemical staining with I-5 monoclonal antibody (mab) (*a, b*) and intracellular injection of HRP (*c*). *a, b*, Camera lucida drawings of different thoracic (T2, T3) and abdominal (A1–A6) segments (in *a*, both sides of T2 and T3; in *b*, only the left side of T3–A6) in the same embryo. Each successive anterior segment is older due to the rostral to caudal developmental gradient. In each segment, the somitic mass (S, surrounding the coelom, C; see A6 segment) spreads in several directions, including ventromedially to form extrinsic limb muscles, ventral longitudinal muscles and ventral transverse muscles. On either side of the midline the mesoderm spreads medially as two distinct masses of small cells which remain separate at their anterior and posterior boundaries and do not fuse: the anterior transverse muscle anlage (TMA) and the posterior longitudinal muscle anlage (LMA). The TMA spreads towards the midline and fuses with the homologous contralateral mass (see A1 segment in *b*); the LMA spreads mainly anteriorly and posteriorly and does not meet or fuse with its homologous contralateral mass. Within the mass of small mesodermal cells comprising the LMA, a single cell, called LM1, near the posterior margin enlarges and begins to stain with the I-5 mab and to extend processes. Shortly thereafter, a single cell, called TM1, appears in the TMA and also begins to stain with the I-5 mab and extend processes. A third large mesodermal cell (LM2) then appears and later other large mesodermal cells appear. *c*, Intracellular injection of HRP into the LM1 cell (at about the stage shown in the T3 segment in *b*) reveals that its growing tips closely resemble neuronal growth cones with many long thin filopodia.

Gastrulation in insects transforms the single layer of the blastoderm into two germ layers, the presumptive mesodermal cells migrating through the gastral groove to become the inner or dorsal layer⁷. The outer (ventral) layer of ectoderm gives rise to epidermis and nervous system; the inner (dorsal) layer of mesoderm gives rise to muscles and some other body parts. The inner layer spreads laterally and, as segmentation proceeds in the ectoderm, cleaves into segmental units of mesodermal cells. The thick lateral segmented masses of mesodermal cells develop into paired, hollow somites surrounding a coelomic cavity (Fig. 1*b*). Each somitic mass spreads ventrolaterally to form the intrinsic muscles of the limb bud, dorsolaterally to form dorsal body muscles, and ventromedially to form extrinsic limb muscles, ventral longitudinal muscles and ventral transverse muscles, as described below.

The mid-ventral region of the ectoderm becomes the neuroepithelium and gives rise to the CNS. A conspicuous basement membrane appears on the dorsal (inner) surface of this ventral ectoderm; this membrane separates the ectoderm from the mesoderm and is mainly secreted by the ectoderm⁸. During the same developmental period in which neuronal growth cones extend along the ventral (or neural) surface of this basement membrane^{5,8,9}, mesodermal cells spread along the dorsal (or mesodermal) surface of the same basement membrane. On either side of each segment, the mesoderm spreads medially as two distinct masses of small cells which remain separate at their boundary and do not fuse—the anterior transverse muscle anlage (TMA) and the posterior longitudinal muscle anlage (LMA) (Fig. 1*a, b*).

Within the mass of small mesodermal cells comprising the LMA, a single cell near the posterior margin enlarges and begins to stain with the I-5 mab. We observe the development of this identified mesodermal cell, called LM1, in the same segment in successively older embryos, or in the same embryo in successively anterior segments (due to the rostral/caudal segmental gradient), as shown in Fig. 1*b*. Shortly after the appearance of LM1, a single cell near the anterior margin of the other mass (the TMA) enlarges and begins to stain with the I-5 mab; this identified cell is called TM1. These two large mesodermal cells (LM1 and TM1) extend processes along the mesodermal surface of the basement membrane in either the longitudinal (LM1) or transverse directions (TM1) (Figs 1, 2, 3*a*). The posteriorly directed growth cone of the LM1 crosses the segment boundary and stops at the homologous location in the next posterior segment to where its anteriorly directed process stops in its own segment; thus, its processes span the distance of one segment. The medially directed process of the

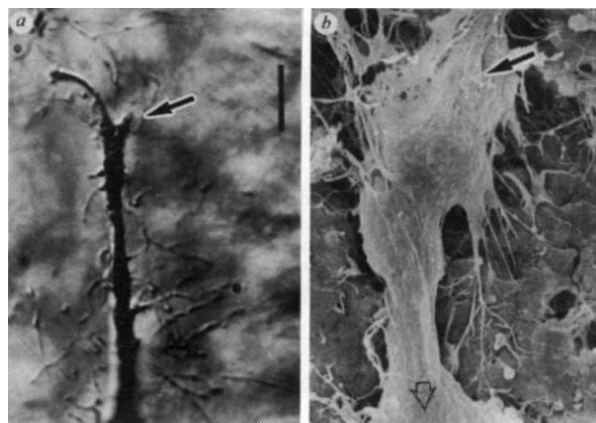


Fig. 2 Photographs of mesodermal growth cones of muscle pioneer LM1 in the body wall of the grasshopper embryo as revealed by intracellular injection of HRP (*a*) and scanning electron microscopy (SEM) (*b*). The black arrow marks the growth cone; the open arrow marks the cell body. Note the many long thin filopodia, which in SEM are seen in contact with the mesodermal surface of the dorsal basement membrane (see text). Scale bars: *a*, 10 μm; *b*, 5 μm.

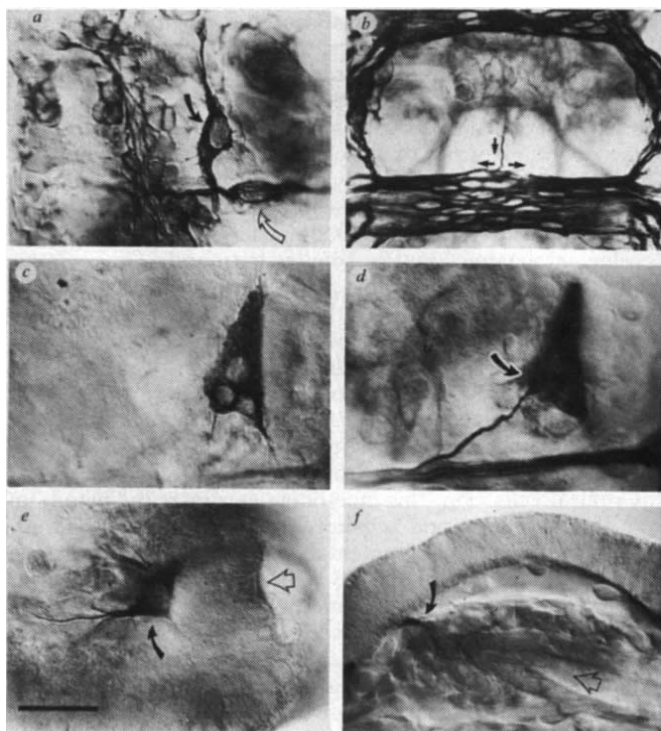


Fig. 3 Muscle pioneers and muscles in the grasshopper embryo. Photographs of mesoderm on the dorsal surface of the neuroepithelium (*a,b*) and in limb buds (*c-f*) in whole mount preparations showing staining of muscle pioneers and other cells with I-5 mab. *a*, The first two muscle pioneers in each body segment of the embryo (see Fig. 1*a,b*) are located on the mesodermal surface of the dorsal basement membrane; both cells have contralateral homologues and are segmentally repeated. Dark arrow: the first longitudinal muscle pioneer, LM1. Open arrow: the first transverse muscle pioneer, TM1. Other staining cells at the left are early differentiating neurones on the other (neural) side of the basement membrane. *b*, The first transverse muscle (the ventral diaphragm muscle) develops along the posterior edge of the TM1 muscle pioneers. The developing axonal pathways of the CNS are seen out of focus below the developing muscles. The arrows mark the median nerve which bifurcates to form the spiracular (transverse) nerves. The spiracular nerve is pioneered by growth cones from respiratory motoneurons using as their substratum the TM1 muscle pioneers. *c,d*, Two different focal planes showing the muscle pioneer for coxal muscle 133 (which originates with a single nucleus and at this stage has three nuclei) and the single growth cone of the identified motoneurone D_4 (arrow, see text; E.E.B., R.K.H. and C.S.G., in preparation) that has left nerve 5 and grown towards and stopped at the muscle pioneer, several days before the differentiation of muscle fibres. *e*, The first muscle pioneer for the extensor tibiae (ETi) muscle in the femur of the limb bud (dark arrow), here shown tightly adhering to the ectodermal epithelium at the site where the extensor tendon invagination has just begun (open arrow). *f*, The same ETi muscle pioneer viewed laterally (dark arrow) at a later stage showing the more fully developed tendon (open arrow) and the connection of the muscle pioneer with the tendon and limb bud epidermis. Note the mass of small mesodermal cells clustered around the tendon. Scale bars: *a,f*, 50 μ m; *b*, 100 μ m; *c,d*, 35 μ m; *e*, 40 μ m.

TM1 stops at the midline (where it meets the process of its contralateral homologue, Fig. 1*a*) while its laterally directed growth cone extends to the lateral margin of the epidermis. In time, more large mesodermal cells appear in the ventromedial mesoderm and extend long processes (for example, LM2, Fig. 1*a,b*). Furthermore, large mesodermal cells that similarly stain with the I-5 mab appear in the ventrolateral mesoderm in the appendages, and in the dorsolateral mesoderm in the dorsal body wall.

Intracellular injections of HRP into these large mesodermal cells reveal that their growing tips closely resemble neuronal growth cones, with many long thin filopodia; thus, we refer to

them as mesodermal growth cones (Figs 1*c*, 2*a*). SEM of their growing processes confirms this observation of profuse filopodia in contact with the basement membrane (Fig. 2*b*). Intracellular injections of LY (450 molecular weight) into these large mesodermal cells (and subsequent staining with anti-LY) reveal that they become dye coupled to the other large mesodermal cells with which they make contact in their own and neighbouring segments. Furthermore, they are dye coupled to many of the smaller mesodermal cells within their same mass.

What is the developmental role of these large mesodermal cells? We call them muscle pioneers, and the following summary of their properties may clarify their role. (1) They appear in the mesoderm early in development when the terrain is relatively simple and distances short, and long before the appearance of muscle fibres or individual muscles (Figs 1*a,b*; 3*a,e*). (2) They have different adhesive properties from other mesodermal cells in that they often flatten and adhere to the ectodermal epithelium (or its overlying basement membrane) (Fig. 3*e,f*). (3) As described above, they extend large growth cones across the inner surface of the ectoderm (or its basement membrane) in both the body wall and appendages. These growth cones stop and insert into the ectoderm at locations which seem to delineate the future muscle insertion sites. (4) By their insertion at the future site of ectodermal invagination, they appear to be intimately involved in the morphogenesis of tendons (in insects, tendons arise as ectodermal invaginations, also called apodemes). For example, the extensor tibiae muscle pioneer in the limb bud inserts at the location of the tendon invagination (Fig. 3*e,f*). (5) Other smaller mesodermal cells cluster around the muscle pioneers and later in development (when the terrain is more complex) differentiate into muscle fibres, using the long processes and ectodermal insertions of the muscle pioneers as a scaffold for their development (Fig. 3*b*). For example, the transverse (or ventral diaphragm) muscle develops along the posterior edge of the TM1 muscle pioneer (Fig. 3*a,b*). Thus, the development of every muscle in the body wall and limb buds thus far examined in the grasshopper embryo is preceded by the earlier appearance of a muscle pioneer which spans the distance and inserts where the future muscle will later appear.

Motoneurone growth cones seem to use the muscle pioneers as one of their guidance cues. The muscle pioneers may have two different roles as they guide growth cones off the peripheral axonal pathways. First, in several cases they serve as a substratum for the navigation of identified growth cones to more distal muscle pioneers. For example, the transverse nerve is pioneered by the growth cones of respiratory motoneurons (C.S.G. and C. M. Bate, unpublished results) which use as their substratum the TM1 muscle pioneer (which pioneers the transverse muscle). In the limb bud, a distal portion of nerve 5 is pioneered by the growth cones of tarsal motoneurons which use as their substratum the retractor unguis muscle pioneers (and distal tendon) in the femur and tibia.

Second, in other cases the muscle pioneers seem to serve as cell-specific signals that guide and stop growth cone extension. For example, in the adult grasshopper, coxal muscle 133 is innervated by two identified excitatory motoneurons (called D_4 and D_5 by morphological homology with the cockroach motoneurons; E.E.B., R.K.H. and C.S.G., in preparation) and one inhibitory motoneurone (CI, common inhibitor). The muscle pioneer for coxal muscle 133 is located within 20 μ m of nerve 5 in the coxa and extends profuse processes as far as the nerve 5 axonal bundle (Fig. 3*c,d*). Most of the growth cones of the leg motoneurons which innervate tibial and tarsal muscles grow out along nerve 5 and extend past this coxal muscle pioneer without leaving the axonal bundle and growing towards the cell. However, the growth cone of motoneurone D_6 , and shortly thereafter the growth cone of motoneurone D_7 , do leave the axonal pathway and stop adjacent to this muscle pioneer. By 45% of development, when growth cones of most of the leg motoneurons have extended past the coxal muscle pioneer, no incorrect growth cones have ever been observed

to innervate this muscle pioneer. This specific association of motoneurone growth cones and muscle pioneer occurs several days before the appearance of differentiated muscle fibres.

There are two important conclusions: muscle pioneers appear to erect a scaffold early in development for later developing muscles, and also to provide guidance cues for motoneurone growth cones. The observation of specific affinities between motoneurone growth cones and muscle pioneers suggests that the muscle pioneers may have differentially labelled surfaces, and that different motoneurone growth cones may be able to distinguish between them. The discovery of muscle pioneers in the grasshopper embryo thus has major implications for the understanding of muscle development and neuromuscular specificity, and offers us the potential for cellular and molecular manipulations in this system.

We thank Susannah Chang and Paul Taghert for their antibodies, Michael Bastiani for help with the SEM, and John Raper for criticism of the manuscript. This work was supported by grants from the McKnight Foundation, NIH and NSF to C.S.G. who is a Sloan Fellow, and the Australian National University to E.E.B. R.K.H. is supported by an NIH Dev. Biol. Training Grant.

Received 9 August; accepted 2 November 1982.

- Landmesser, L. *J. Physiol., Lond.* **284**, 371 (1978).
- Dennis, M. J. A. *Rev. Neurosci.* **4**, 43 (1981).
- Lance-Jones, C. & Landmesser, L. *Proc. R. Soc. B214*, 1 (1981).
- Chang, S., Ho, R. K. & Goodman, C. S. *Dev. Biol.* (submitted).
- Taghert, P. T., Bastiani, M., Ho, R. K. & Goodman, C. S. *Dev. Biol.* (in the press).
- Ho, R. K. & Goodman, C. S. *Nature* **297**, 404 (1982).
- Anderson, D. T. in *Developmental Systems: Insects* Vol. 1 (eds Counce, S. J. & Waddington, C. H.) Academic, New York, 1972).
- Bate, C. M. & Grunewald, E. B. *J. Embryol. exp. Morph.* **61**, 317 (1981).
- Goodman, C. S., Raper, J. A., Ho, R. K. & Chang, S. *40th Symp. Soc. dev. Biol.*, 275 (1982).

A dopamine- and cyclic AMP-regulated phosphoprotein enriched in dopamine-innervated brain regions

S. Ivar Walaas, Dana W. Aswad* & Paul Greengard

Departments of Pharmacology and Psychiatry, Yale University School of Medicine, New Haven, Connecticut 06510, USA

Several mammalian neurotransmitter candidates, for example, serotonin¹, dopamine^{2,3} and noradrenaline^{4,5}, may exert some of their synaptic effects by regulating protein phosphorylation systems^{6,7}. Comparison of the regional distribution of brain phosphoproteins with neurotransmitter systems may help to identify the specific phosphoproteins involved in the functions of particular neurotransmitters. Here we report the association of one such phosphoprotein with the dopamine pathways in brain. This protein, of apparent molecular weight (MW) 32,000 (32K), seems to be present only in nervous tissue. Its regional distribution within the brain is very similar to the pattern of dopamine-containing nerve terminals; more specifically, the protein appears to be enriched in those dopaminergic neurones which possess D-1 receptors (dopamine receptors coupled to adenylate cyclase⁸). The state of phosphorylation of the protein in these dopaminergic neurones can be regulated by both dopamine and cyclic AMP. These results suggest that the phosphoprotein may mediate certain of the trans-synaptic effects of dopamine acting on dopaminergic neurones.

Endogenous phosphorylation of brain proteins was examined in total homogenates from rat neocortex and neostriatum (caudatoputamen). Incubation of these preparations with [γ -³²P]ATP and Mg²⁺, in the absence or presence of cyclic AMP, followed by electrophoretic separation of the phosphorylated

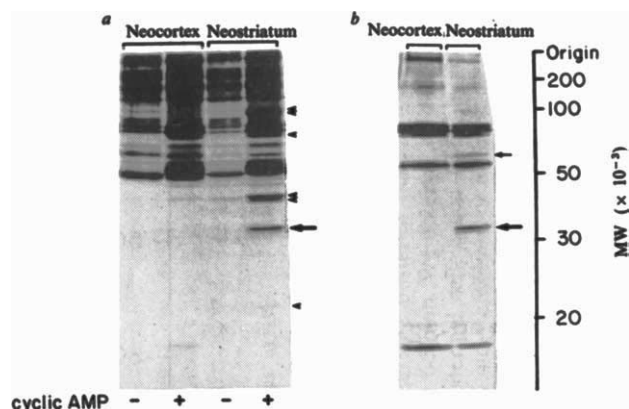


Fig. 1 Demonstration of the 32K phosphoprotein (DARPP-32) in neostriatum. Presence of DARPP-32 in homogenates (a) and acid extracts (b). The heavy arrow in a and b indicates DARPP-32. Arrowheads in a indicate other proteins phosphorylated in the presence of cyclic AMP that are present in the neostriatum, but not in the neocortex. The light arrow in b indicates a 62K phosphoprotein.

Methods: a, Male Sprague-Dawley rats (200 g) were decapitated, the brain removed and the neostriatum and neocortex dissected and homogenized in ice-cold standard buffer containing 10 mM Tris-HCl (pH 7.4), 2 mM EDTA and 0.1 mM of the protease inhibitor phenylmethylsulphonyl fluoride. Aliquots containing 0.2 mg protein were added to a reaction mixture (100 μ l final volume) containing (final concentration): 10 mM Tris-HCl pH 7.4, 1 mM EDTA, 1 mM EGTA, 1 mM dithiothreitol, 6 mM Mg²⁺ and 20 μ M [γ -³²P]ATP (specific activity 20–30 Ci mmol⁻¹), in the absence (-) or presence (+) of 2 μ M cyclic AMP plus 1 mM isobutylmethylxanthine. The samples were preincubated at 30 °C for 1 min. The reaction was initiated by adding [γ -³²P]ATP, and terminated after 10 s by the addition of SDS-containing 'stop solution' and immediate boiling for 2 min. Samples containing 100 μ g of protein were then subjected to SDS-polyacrylamide gel electrophoresis and autoradiography as described elsewhere⁹. The autoradiogram obtained in a typical experiment is shown. b, Aliquots (0.1 ml) of the neocortex and neostriatum homogenates containing 0.5 mg protein were incubated at 37 °C for 15 min to dephosphorylate the phosphoproteins²⁴, added to 5 ml of ice-cold Zn-acetate (5 mM), and centrifuged at 4,000g for 15 min. The pellets were resuspended in 0.5 ml of ice-cold Na-citrate-phosphate buffer (10 mM, pH 2.8), containing 0.1% (v/v) Triton X-100 and 2 μ g ml⁻¹ pepstatin A. (Inclusion of this concentration of non-ionic detergent was found to improve the reproducibility of the extraction procedure and the following phosphorylation reaction.) The suspension was centrifuged at 28,000g for 15 min, 0.4 ml of the supernatant was adjusted to pH 6 with 0.5 M Na₂HPO₄, centrifuged at 15,000g for 15 min, and the supernatants were used as the final acid extracts. The amount of dephospho-DARPP-32 in 60- μ l aliquots of the final extracts was determined by 'back phosphorylation' (in 100 μ l final volume) in a reaction mixture containing (final concentration): 50 mM HEPES pH 7.4, 10 mM Mg²⁺, 1 mM EDTA, 1 mM EGTA, 3 μ M [γ -³²P]ATP (specific activity 20–40 Ci mmol⁻¹) and 10 nM of the purified catalytic subunit of cyclic AMP-dependent protein kinase prepared according to Beavo *et al.*²⁵. The reaction was carried out at 30 °C for 60 min, allowing maximal incorporation of ³²P into DARPP-32. Phosphorylated samples, containing 20 μ g of protein, were then subjected to SDS-polyacrylamide gel electrophoresis and autoradiography, with the results shown.

proteins and autoradiography of the dried gels⁹, demonstrated a protein with an apparent MW of 32K which was phosphorylated in the presence of cyclic AMP in the neostriatum but not in the cortical preparation (Fig. 1a). We refer to this protein as DARPP-32 (dopamine- and adenosine 3':5'-monophosphate-regulated phosphoprotein). Several other proteins present in the neostriatum that were phosphorylated in the presence of cyclic AMP, but not detected in the neocortex, had apparent molecular weights of 96–98K, 81K, 39–40K and 21K (Fig. 1a). Two major groups of proteins with apparent molecular weights

* Present address: Department of Psychobiology, University of California, Irvine, California 92717, USA.

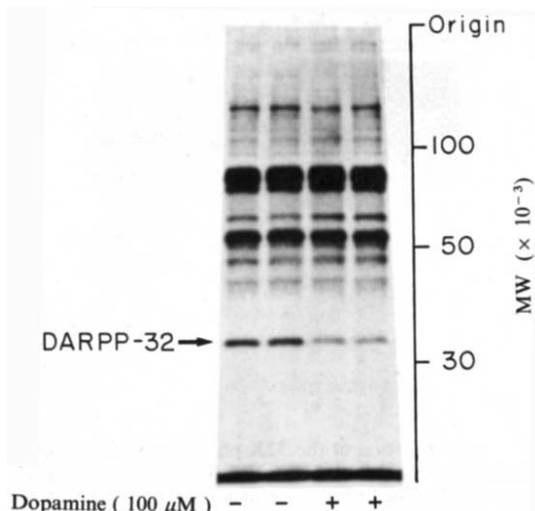


Fig. 2 Regulation by dopamine of protein phosphorylation in intact slices of rat neostriatum. Male Sprague-Dawley rats (200 g) were decapitated, and the brains removed. The neostriatum was dissected and chopped into 0.4 mm coronal slices with a McIlwain tissue chopper. The slices were rapidly suspended in 10 ml Krebs-Ringer bicarbonate buffer (KRB) pH 7.4 at 30 °C (composition in mM: Na⁺, 150; K⁺, 5; Ca²⁺, 1.0; Mg²⁺, 1.5; PO₄³⁻, 1.5; HCO₃⁻, 25; Cl⁻, 130; ascorbate, 1.0; D-glucose, 10). The medium was continuously gassed with H₂O-saturated 95% O₂/5% CO₂ throughout the experiment. Slices were preincubated for 1 h, with one change of medium. Isobutylmethylxanthine was then added (final concentration, 20 μM) and the preincubation continued for 5 min. Dopamine (final concentration 100 μM) was then added to experimental slices. Incubation was carried out for 10 min and was stopped by aspirating the KRB, adding 5 ml of ice-cold Zn-acetate (5 mM) and rapidly homogenizing the tissue. The phosphoproteins were subjected to acid extraction, back phosphorylation, SDS-polyacrylamide gel electrophoresis (40 μg protein per lane), and autoradiography as described in Fig. 1 legend. The autoradiogram obtained in a typical experiment, using duplicate experimental and control tissue samples, is shown. The effect of dopamine and of cyclic AMP on the amount of dephospho-DARPP-32 in the polyacrylamide gels was determined quantitatively, using the procedure described in Table 1 legend. In each of 10 experiments, neostriatal slices were incubated in quadruplicate in the absence or presence of 100 μM dopamine or 4 mM 8-bromo-cyclic AMP; each sample was then assayed in duplicate for dephospho-DARPP-32. The amount of dephospho-DARPP-32 in samples incubated with dopamine and with 8-bromo-cyclic AMP was 78 ± 3% and 75 ± 6% (mean ± s.e.m.) of the amount present in paired control samples, respectively; both values were significantly different from control samples ($P < 0.05$, Wilcoxon's paired comparison test²⁶).

of 80–86K and 50–55K, and some minor phosphoproteins, exhibited cyclic AMP-dependent phosphorylation in both brain regions (Fig. 1a).

In separate experiments, we found that DARPP-32, but not the other prominent phosphoproteins found only in the neostriatum, could be quantitatively extracted from neostriatal homogenates by a modification of the acid extraction technique originally developed^{9,10} for synapsin I (protein I¹¹). This technique allowed a quantitation of the dephospho-forms of DARPP-32 and of synapsin I by phosphorylation of the extracts with [γ -³²P]ATP in the presence of the catalytic subunit of cyclic AMP-dependent protein kinase¹⁰. Therefore, from the group of proteins found in the neostriatum but not in the neocortex, we selected DARPP-32 for further study. Figure 1b shows the presence of both DARPP-32 and a minor 62K phosphoprotein in acid extracts of the neostriatum but not in extracts of the neocortex. In contrast, the 80–86K doublet representing synapsin I was found in approximately equal concentration in acid extracts from both brain regions (Fig. 1b).

A more detailed distribution of DARPP-32 is given in Table 1. DARPP-32 was found to be highly enriched in the substantia

Table 1 Regional distribution of DARPP-32 and synapsin I in selected regions of the rat central nervous system

Region	DARPP-32 (pmol ³² P incorporated per mg tissue protein)	Synapsin I
Substantia nigra	15.50 ± 1.90	27.50 ± 3.10
Neostriatum	11.30 ± 1.00	34.70 ± 4.80
Olfactory tubercle	7.92 ± 1.00	24.33 ± 2.55
Nucleus accumbens	4.75 ± 0.80	25.61 ± 5.28
Neocortex	1.83 ± 0.11	32.00 ± 2.93
Septum	0.65 ± 0.06	25.80 ± 2.16
Spinal cord	0.55 ± 0.05	17.21 ± 2.02
Ventral tegmental area	0.17 ± 0.04	24.32 ± 5.88

Male Sprague-Dawley rats (150–200 g) were decapitated, and their brains and spinal cords were removed and placed in ice-cold standard buffer. The brain regions were dissected from coronal slices (0.6–0.8 mm thick) prepared with a McIlwain tissue chopper, following the atlas of König and Klippel²¹. The neocortex was taken from two slices at levels A 6,500–A 8,000, and did not include the piriform cortex. The ventral tegmental area included the interpeduncular nucleus and the A10 region²², while the substantia nigra included both the pars reticulata and pars compacta. The samples were homogenized (5% w/v) in ice-cold standard buffer and the protein content analysed²³. Aliquots containing 1 mg protein were acid-extracted and phosphorylated as described in Fig. 1 legend. The phosphoproteins were separated by SDS-polyacrylamide gel electrophoresis, autoradiographed and excised from the dried gels using the autoradiograms as guides. ³²P content was determined by liquid scintillation counting of the gel pieces. Data represent the mean ± s.e.m. of four samples, each phosphorylated in duplicate.

nigra and neostriatum, while slightly lower concentrations of the protein were seen in the olfactory tubercle and the nucleus accumbens. The neocortex contained less than 15% of the amount of DARPP-32 found in the substantia nigra (Table 1). Only trace amounts were found in the septum, spinal cord and ventral tegmental area. In contrast, synapsin I, which is present in almost all nerve terminals¹², was relatively evenly distributed among the various regions examined. Peptide maps of DARPP-32, obtained using *Staphylococcus aureus* V8 protease^{13,14}, indicated that the 32K phosphoproteins seen in all the brain regions examined were the same (not shown). Moreover, these peptide maps of DARPP-32 were markedly different from those of any other brain phosphoprotein characterized to date. Neither DARPP-32 nor synapsin I was detected (< 0.05 pmol ³²P incorporated per mg tissue protein) in extracts from any of several non-neuronal tissues examined, including liver, heart, lung, spleen, kidney and skeletal muscle.

The highly specific regional distribution of DARPP-32 suggested that the phosphoprotein might be associated with the dopaminergic system in the brain, as DARPP-32 was found in high concentrations only in those regions having a high density of dopaminergic neurones or their processes¹⁵. Moreover, this phosphoprotein was enriched in regions that contain the D-1 class of dopamine receptor, for example, the substantia nigra^{16,17}, neostriatum¹⁸, nucleus accumbens¹⁹ and olfactory tubercle¹⁹, but not in a region containing only the D-2 class of dopamine receptor—the ventral tegmental area^{16,20}. These results led us to examine the effect of dopamine and cyclic AMP on the state of phosphorylation of DARPP-32. Incubation of neostriatal slices with 100 μM dopamine for 10 min induced a significant decrease in the amount of dephospho-DARPP-32 (Fig. 2). Similar results were observed after incubation with 4 mM 8-bromo-cyclic AMP for 10 min. The total amount of DARPP-32 was unchanged during these incubations (data not shown), indicating that dopamine and cyclic AMP had caused the conversion of DARPP-32 from the dephosphorylated form to the phosphorylated form in the neostriatum. In one series of experiments, we determined that the amount of DARPP-32 in the phosphorylated form in neostriatal slices incubated in the absence or presence of dopamine represented ~20% and 40%, respectively, of the total DARPP-32 present

(S.I.W. and P.G., in preparation). Thus, incubation with dopamine caused an approximate doubling of the amount of DARPP-32 in the phosphorylated form in the experimental conditions used. Consistent with the notion that dopamine stimulates the phosphorylation of DARPP-32 by activating dopamine receptors and increasing intracellular cyclic AMP levels, we have found that fluphenazine, a potent antipsychotic dopamine receptor blocker⁸, reversed the effect of dopamine, while addition of other neurotransmitters known to stimulate cyclic AMP synthesis in brain tissue (for example, noradrenaline and serotonin^{4,5}) induced no detectable change in the state of DARPP-32 phosphorylation. Moreover, a 30-s depolarization-induced calcium influx into neostriatal neurones—a procedure expected to activate calcium-dependent protein kinases¹⁴—had no effect on the amount of dephospho-DARPP-32 (not shown).

Thus, the phosphoprotein DARPP-32 described here may be functionally related to dopaminergic neurotransmission. The

finding that cyclic AMP can mimic the effect of dopamine on the phosphorylation of DARPP-32 suggests that this effect of dopamine is mediated through D-1 receptors located on dopaminergic neurones. Indeed, this interpretation agrees with the detailed anatomical distribution (S.I.W. and P.G., in preparation) of this protein. Further elucidation of the characteristics of DARPP-32 should help to clarify molecular basis of dopamine effects mediated through the D-1 class of dopamine receptors.

The catalytic subunit of cyclic AMP-dependent protein kinase was given by Dr Angus C. Nairn of this laboratory. This work was supported by USPHS grants MH-17387 and NS-08440 and a grant from the McKnight Foundation (to P.G.). Part of the work was done while S.I.W. was on leave of absence from the Norwegian Defense Research Establishment, supported by a NATO Science Fellowship from the Norwegian Council for Scientific and Industrial Research.

Received 15 September; accepted 1 November 1982.

1. Dolphin, A. C. & Greengard, P. *Nature* **289**, 76–78 (1981).
2. Nestler, E. J. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7479–7483 (1980).
3. Tsou, K. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6075–6079 (1982).
4. Nathanson, J. A. *Physiol. Rev.* **57**, 157–256 (1977).
5. Bloom, F. E. *Rev. Physiol. Biochem. Pharmac.* **74**, 1–103 (1975).
6. Greengard, P. *Fedn Proc.* **38**, 2208–2217 (1979).
7. Greengard, P. *Harvey Lect.* **75**, 277–331 (1981).
8. Keibabian, J. W. & Calne, D. B. *Nature* **277**, 93–96 (1979).
9. Ueda, T. & Greengard, P. *J. biol. Chem.* **252**, 5155–5163 (1977).
10. Forn, J. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5195–5199 (1978).
11. Huttner, W. B., De Camilli, P., Schiebler, W. & Greengard, P. *Soc. Neurosci. Abstr.* **7**, 441 (1981).
12. De Camilli, P., Cameron, R. & Greengard, P. *J. Cell Biol.* **87**, 72a (1980).
13. Cleveland, D. W., Fischer, S. G., Kirschner, M. W. & Laemmli, U. K. *J. biol. Chem.* **252**, 1102–1106 (1977).

14. Huttner, W. B. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5402–5406 (1979).
15. Lindvall, O. & Björklund, A. *Handbook of Psychopharmacology* (eds Iversen, L. L., Iversen, S. D. & Snyder, S. H.) 139–231 (Plenum, New York, 1978).
16. Premont, J. et al. *FEBS Lett.* **68**, 99–104 (1976).
17. Keibabian, J. W. & Saavedra, J. M. *Science* **193**, 683–685 (1976).
18. Keibabian, J. W., Petzold, G. L. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **69**, 2145–2149 (1972).
19. Horn, A. C., Cuello, A. C. & Miller, R. J. *J. Neurochem.* **22**, 265–270 (1974).
20. Gundlach, A. L., McDonald, D. & Beart, P. M. *J. Neurochem.* **39**, 890–894 (1982).
21. König, J. F. R. & Klippel, R. A. *The Rat Brain* (Williams & Wilkins, Baltimore, 1963).
22. Fonnum, F., Walaas, I. & Iversen, E. *J. Neurochem.* **29**, 221–230 (1977).
23. Peterson, G. *Analyt. Biochem.* **83**, 346–356 (1977).
24. Dolphin, A. C. & Greengard, P. *J. Neurochem.* **36**, 1627–1631 (1981).
25. Beavo, J. A., Bechtel, P. J. & Krebs, E. G. *Meth. Enzym.* **38C**, 299–309 (1974).
26. Hodges, J. L. Jr & Lehmann, E. L. *Basic Concepts of Probability and Statistics* (Holden-Day, San Francisco, 1963).

Developmental stability and enzyme heterozygosity in rainbow trout

Robb F. Leary, Fred W. Allendorf
& Kathy L. Knudsen

Department of Zoology, University of Montana, Missoula,
Montana 59812, USA

The developmental pathways of organisms are genetically adjusted to produce the characteristic morphology of the species regardless of variations in internal and external conditions during development. This 'developmental buffering', however, is not always precise. Bilateral characters of an organism are often asymmetric, that is, different in size, shape or number. Fluctuating asymmetry occurs when the difference between a character on the left and right sides of individuals is normally distributed about a mean of zero¹. This type of asymmetry results from the inability of an organism to develop precisely along determined paths^{1,2} and can be used as a measure of developmental stability^{1,3–5}. Increased developmental stability would be reflected by reduced amounts of fluctuating asymmetry. We have now examined the relationship between the amount of fluctuating asymmetry for five bilateral characters and heterozygosity at 13 polymorphic loci in a population of rainbow trout (*Salmo gairdneri*). Our results indicate a significant negative correlation between the proportion of heterozygous loci and the proportion of asymmetric characters. These data provide evidence that individuals with greater heterozygosity within a population have increased developmental stability.

Lerner⁶ has argued that more heterozygous individuals should be characterized by increased developmental stability. Evidence supporting this hypothesis, however, has been indirect. Morphological variation is often greater both within and between crosses from inbred strains than between crosses

of two different inbred strains^{6,7}. More heterozygous populations of a lizard (*Uta stansburiana*)⁸, two freshwater bivalves (*Elliptio complanata* and *Lampsilis radiata*)⁹ and a fish (*Poeciliopsis monacha*)¹⁰, have lower amounts of fluctuating asymmetry. The average coefficient of variation for seven meristic traits was lower among heterozygotes than among homozygotes at five loci in the killifish, *Fundulus heteroclitus*¹¹. The variance of two morphometric characters among heterozygotes was lower than among homozygotes at six loci in the monarch butterfly, *Danaus plexippus*¹². We are not aware, however, of any studies showing a direct association between heterozygosity and developmental stability between individuals within populations⁸.

For 50 rainbow trout (26 males and 24 females) from the Arlee strain at the Jocko River State Trout Hatchery, Montana, we determined the genotypes at 40 loci with horizontal starch gel electrophoresis¹³ and experimental matings. This strain has been maintained in isolation at this hatchery for at least 12 generations (30 yr) and has a slightly greater than average amount of genetic variation compared with other natural and hatchery populations of rainbow trout¹⁴ (average heterozygosity per locus = 0.071). The following enzymes and loci were polymorphic: creatine kinase (*Ck1*), glycerol-3-phosphate dehydrogenase (*G3p1*), isocitrate dehydrogenase (*Idh2* and *Idh3*, 4), lactate dehydrogenase (*Ldh4*), malate dehydrogenase (*Mdh3,4*), malic enzyme (*Me1*), phosphoglucosmutase (*Pgm1-t* and *Pgm2*), sorbitol dehydrogenase (*Sdh*), and superoxide dismutase (*Sod*).

All salmonid fishes, including rainbow trout, are derived from a tetraploid ancestor of some 50–100 Myr years ago¹⁵. Some of the enzyme loci in rainbow trout are still functionally duplicated, that is, the two loci show no evidence of structural or regulatory divergence¹⁶. The two pairs of such duplicated loci (*Idh3,4* and *Mdh3,4*) in this study were treated as a single functional locus. Thus, individuals having only a single electrophoretic form for both loci were considered 'homozygous' and individuals showing two or more electrophoretic forms were considered 'heterozygous'.

Table 1 Number of heterozygous loci and asymmetric characters for 50 individuals in a population of rainbow trout

No. of asymmetric characters	No. of heterozygous loci					
	0	1	2	3	4	5
0	—	—	—	2	1	2
1	2	2	4	4	1	—
2	1	3	10	6	2	—
3	—	3	2	3	—	—
4	—	—	1	—	—	—
5	1	—	—	—	—	—
Mean	2.25	2.13	2.00	1.67	1.25	0.00

The *Pgm1-t* locus is a temporal regulatory locus¹⁷ that affects the expression of the *Pgm1* locus in liver¹⁸. There is usually no detectable PGM1 enzyme in the liver of rainbow trout. However, a variant allele at *Pgm1-t* with additive inheritance results in the expression of this locus in liver. The genotypes of all the fish in our sample with PGM1 in the liver were determined to be heterozygous by experimental matings¹⁸.

The counts of five bilateral meristic characters were taken on the left and right side of each fish: gill rakers on the lower first branchial arch, gill rakers on the upper first branchial arch, mandibular pores, rays in the pectoral fins, and rays in the pelvic fins. These characters were chosen because they could be easily and accurately counted. Pairwise *F* and *t*-tests revealed no significant differences in the variance and means of the distributions of these characters between the sexes. Therefore, we combined the data from the sexes in our analyses. There is no significant correlation between the number of gill rakers on the upper and lower branchial arches. These are apparently independently determined characters. There is no tendency for a character to have a higher count on one side of the organism than the other. The modal value of the distribution of the differences of each character is zero. Furthermore, there are no significant differences between the variances and means of the distribution of these characters on the left and right sides. All of the characters, therefore, exhibit fluctuating asymmetry.

The null hypothesis is that there is no correlation between the proportion of loci that are heterozygous and the proportion of traits that are asymmetric within individuals. The alternative hypothesis is that there is a negative correlation between the proportion of heterozygous loci and the proportion of asymmetric characters. The number of heterozygous loci and the number of asymmetric characters per fish both ranged from zero to five (Table 1). The correlation between these proportions is significant using both a *t*-test with an arc sine transformation for proportions¹⁹ ($r = -0.40$, $t = -3.07$, $P < 0.002$) and the non-parametric Spearman's rank correlation test ($r = -0.30$, $t = -2.18$, $P < 0.02$).

We also examined the relationship between the number of heterozygous loci and the magnitude of the asymmetry per individual. The magnitude of the asymmetry per individual was computed as the sum of the absolute values of the differences of the five bilateral characters. We did not correct the absolute values for scaling⁵ as the mean counts of each character were not markedly different (range 7.9–14.4), and there was no association between the number of fish asymmetric for a given character and the character mean ($r = -0.161$, $t = -0.283$, $P > 0.50$). The Spearman's rank correlation between the number of heterozygous loci and the magnitude of asymmetry is significant ($r = -0.31$, $t = -2.19$, $P < 0.02$).

We next tested whether this relationship between heterozygosity and less asymmetry was associated with certain individual loci or with a general heterozygous effect. We examined the amount of asymmetry for homozygotes and heterozygotes at all loci in which there were over five heterozygotes (Table 2). Heterozygotes have a smaller mean proportion of asymmetric traits at seven of the eight loci. The heterozygotes at *Idh3,4* ($t =$

Table 2 Mean percentage of asymmetric characters for homozygotes and heterozygotes at eight loci in rainbow trout

Locus	Homozygotes	Heterozygotes	Difference
<i>Idh2</i>	36.7	29.1	+7.6
<i>Idh3,4</i>	42.7	32.5	+10.2*
<i>Ldh4</i>	34.7	41.3	-6.6
<i>Mdh3,4</i>	36.7	33.3	+3.4
<i>Pgm1-t</i>	38.2	27.9	+10.3*
<i>Pgm2</i>	35.8	32.5	+3.3
<i>Sdh</i>	36.2	30.7	+5.5
<i>Sod</i>	35.7	34.2	+1.5

* $P < 0.05$.

-2.04 , $P < 0.03$) and *Pgm1-t* ($t = -2.09$, $P < 0.03$) have a significantly lower mean proportion of asymmetric characters than the homozygotes. There was no significant difference in the mean proportion of asymmetric characters between the homozygotes and heterozygotes at the other loci. We eliminated *Idh3,4* and *Pgm1-t* from the data to determine if the overall effect is due only to heterozygosity at these two loci. The correlation between the proportion of heterozygous loci per fish and the proportion of asymmetric characters remains significant ($r = -0.30$, $t = -2.16$, $P < 0.03$).

These results demonstrate that heterozygosity in general and heterozygosity at specific loci are associated with increased developmental stability in this population of rainbow trout. This observation supports Lerner's hypothesis that more heterozygous individuals within a population have increased developmental stability. Nevertheless, it is not known whether these effects are due to the action of the loci examined or to chromosomal segments that are marked by these loci.

The counts of meristic characters in fish are determined early during development^{20–22}. During this period, the liver is responsible for the release of energy stored in the yolk^{23,24}. *Idh3,4* encodes the only IDH isozymes present in liver. Thus, the loci showing significant single locus effects produce liver enzymes involved in the two major biochemical pathways producing the energy for metabolic activities during development²⁴: glycolysis (PGM) and the citric acid cycle (IDH). Heterozygosity at these loci may result in a more constant and efficient flux through glycolysis and the citric acid cycle, resulting in increased developmental stability. Therefore, there is a possible specific mechanism to explain the observed relationship between heterozygosity and developmental stability that we intend to test in future experiments.

This work was supported by NSF grants DEB-8004681 and ISP-8011449. We thank H. McPherson and N. Ryman for stimulating conversation.

Received 2 August; accepted 13 October 1982.

- Van Valen, L. *Evolution* **16**, 125–142 (1962).
- Mason, L. G., Ehrlich, P. R. & Emmel, T. C. *Evolution* **21**, 85–91 (1967).
- Thoday, J. M. *Heredity* **12**, 401–415 (1958).
- Soulé, M. *Am. Nat.* **101**, 141–160 (1967).
- Felley, J. *Copeia* **1980**, 18–29 (1980).
- Lerner, I. M. *Genetic Homeostasis* (Oliver and Boyd, Edinburgh, 1954).
- Robertson, F. W. & Reeve, E. C. R. *Nature* **170**, 286 (1952).
- Soulé, M. *Evolution* **33**, 396–401 (1979).
- Kat, P. W. *Am. Nat.* **119**, 824–832 (1982).
- Vrijenhoek, R. G., & Lerman, S. *Evolution* **36**, 768–776 (1982).
- Mitton, J. B. *Nature* **273**, 661–662 (1978).
- Eanes, W. F. *Nature* **276**, 263–264 (1978).
- Allendorf, F. W., Mitchell, N. J., Ryman, N. & Ståhl, G. *Heredity* **86**, 179–190 (1977).
- Allendorf, F. W. & Phelps, S. R. *Ecol. Bull.* **34**, 37–52 (1981).
- Ohno, S. *Animal Cytogenetics 4* (Gebrüder Borntraeger, Berlin, 1974).
- Bailey, G. S., Wilson, A. C., Halver, J. C. & Johnson, C. L. *J. biol. Chem.* **245**, 5927–5940 (1970).
- Paigen, K. in *Physiological Genetics* (ed. Scandalios, J. G.) 1–61 (Academic, New York, 1979).
- Allendorf, F. W., Knudsen, K. L. & Phelps, S. R. *Genetics* **102**, 259–268 (1982).
- Sokal, R. R. & Rohlf, F. J. *Biometry*, 427–428 (Freeman, San Francisco, 1981).
- Barlow, G. W. *Syst. Zool.* **10**, 105–117 (1961).
- Garside, E. T. *J. Fish. Res. Bd Can.* **23**, 1537–1551 (1966).
- Ali, M. Y. & Lindsey, C. C. *Can. J. Zool.* **52**, 959–976 (1974).
- Terner, C. *Comp. Biochem. Physiol.* **25B**, 989–1003 (1968).
- Boulekbache, H. *Am. Zool.* **21**, 377–389 (1981).

Glucose modulates Mg^{2+} fluxes in pancreatic islet cells

J. C. Henquin, T. Tamagawa, M. Nenquin & M. Cogneau*

Unité de Diabète et Croissance, University of Louvain School of Medicine, UCL 54.74, 1200 Brussels, Belgium

* Unité de Chimie Inorganique et Nucléaire, University of Louvain Faculty of Sciences, 1348 Louvain-la-Neuve, Belgium

Magnesium, the most abundant intracellular divalent cation, is an essential cofactor for many enzyme systems¹, but it remains unknown as to whether variations in the cytoplasmic concentration of ionized Mg^{2+} directly control cellular processes. Experiments with adrenal medullary cells made 'leaky' by exposure to high electric fields provided evidence that Mg^{2+} could influence hormone release not only by competing with Ca^{2+} for entry into the cell, but also at intracellular sites controlling exocytosis^{2,3}. A similar conclusion was reached for insulin release in a study⁴ using isolated rat islets also subjected to high voltage discharges. There is no experimental evidence, however, that physiological stimuli influence Mg^{2+} movements in intact secretory cells. We report here that $^{28}Mg^{2+}$ fluxes in pancreatic islet cells are markedly modified by glucose, the physiological stimulus of insulin release, but not by its non-insulinotropic analogue, 3-*O*-methylglucose.

^{28}Mg ($t_{1/2} = 21.3$ h) was produced⁵ via the nuclear reaction ^{27}Al (α , 3p) $\rightarrow$ ^{28}Mg using high energy α -particles at the isochronous cyclotron at Louvain-la-Neuve. A yield of 8 μCi per μAh was obtained for α -particles of 110 MeV on a target of high purity aluminium (1.3 g cm^{-2} ; Goodfellow, UK). The radionuclide formed in a carrier-free state was purified on cellulose columns and eluted as $MgCl_2$.

Control experiments, done in the presence of a non-stimulatory concentration of glucose (3 mM), have shown that ^{28}Mg uptake by rat islets is linear for ~ 15 min, but is far from complete even after 180 min (data not shown). Long-term maintenance of the islets in tissue culture would be necessary to reach isotopic equilibrium. The 10-min uptake, which may thus be taken as representative of ^{28}Mg influx, was increased by 15 mM glucose (Table 1). This stimulation was much less marked in the presence (17%) than in the absence (47%) of calcium, the omission of which also resulted in an increase (40%; $P < 0.001$) in ^{28}Mg influx in the glucose-free medium. 3-*O*-methylglucose had no effect. ^{28}Mg uptake by islets incubated for 120 min in the absence of glucose increased when the concentration of extracellular Mg^{2+} was increased and was augmented when Ca^{2+} was absent (Table 1). Glucose, but not 3-*O*-methylglucose, stimulated ^{28}Mg uptake in all conditions. The largest (185%) and smallest (18%) relative increases brought about by glucose were observed in the presence of the lowest (0.2 mM) and highest (5 mM) concentrations of Mg^{2+} , respectively.

Further experiments were done to establish that the stimulation of ^{28}Mg uptake by glucose did not merely result from an increase in extracellular binding of Mg^{2+} . Islets were first incubated for 120 min without or with 15 mM glucose, in the presence of 2 mM $CaCl_2$ and 1 mM $^{28}MgCl_2$ (30–45 mCi $mmol^{-1}$). After three rapid rinses at room temperature to remove the radioactive solution, the islets were washed twice for 15 min at 10 °C, with a glucose-free, Ca -, Mg - and Na -deficient medium supplemented with 50 μM EDTA. After such washing, the amount of ^{28}Mg retained by the islets was still much higher when the initial incubation period had been performed in the presence of 15 mM glucose (15.5 ± 0.61 pmol per islet) than in its absence (5.36 ± 0.23 pmol per islet for 15 batches of 6 islets each). Although these values are lower than those measured with the double tracer technique (Table 1), the absolute increase in ^{28}Mg uptake is very similar (10.1 compared with 11.9 pmol

Table 1 Effects of glucose and 3-*O*-methylglucose on $^{28}Mg^{2+}$ uptake by islet cells

Ionic composition of medium (mM)		$^{28}Mg^{2+}$ uptake (pmol per islet)		
Mg^{2+}	Ca^{2+}	Controls	Glucose (15 mM)	3- <i>O</i> -Methylglucose (15 mM)
Incubation for 10 min				
1	2	4.20 ± 0.19	$4.93 \pm 0.28^*$	4.48 ± 0.18
1	0	5.97 ± 0.21	$8.80 \pm 0.31^{\dagger}$	6.03 ± 0.23
Incubation for 120 min				
1	2	11.4 ± 0.49	$23.3 \pm 0.58^{\dagger}$	11.7 ± 0.39
1	0	17.3 ± 0.54	$31.9 \pm 1.15^{\dagger}$	17.0 ± 0.29
0.2	2	2.85 ± 0.12	$8.12 \pm 0.28^{\dagger}$	—
5	2	37.4 ± 1.54	$44.3 \pm 1.82^*$	—

Mg^{2+} uptake was measured relative to that of sucrose, which is restricted to the extracellular space. The techniques used have been described previously^{7,22} and are only outlined below. All experiments were carried out at 37 °C, with a bicarbonate-buffered solution gassed with O_2/CO_2 (94:6) pH 7.4, supplemented with 5 $mg\ ml^{-1}$ bovine serum albumin; its usual ionic composition was (in mM): NaCl 118, KCl 6, $CaCl_2$ 2, $MgCl_2$ 1, $NaHCO_3$ 24. For uptake studies, the medium also contained 10 mM HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulphonic acid) to prevent any change in pH in the small volumes of solutions used in these experiments. Isolated rat islets were obtained by collagenase digestion of the pancreas of fed male Wistar rats. After preliminary incubation for 30 min in the presence of 3 mM glucose, 1 mM $MgCl_2$ and 2 mM $CaCl_2$, batches of six islets were incubated in 50 μl of prewarmed medium layered on silicone oil. The incubation medium contained the indicated concentrations of Mg^{2+} , Ca^{2+} , glucose and 3-*O*-methylglucose and was supplemented with 0.25 mM [$6,6$ - 3H] sucrose (0.1 Ci $mmol^{-1}$; Amersham) and a trace amount of ^{28}Mg (0.6–1.2 μCi per 50 μl). After incubation for 10 or 120 min, the islets were separated from the radioactive solution by centrifugation through the oil and the radioactivity entrapped in the tissue was counted by liquid scintillation. Blanks without islets did not differ from the background. Values are mean $\pm$ s.e.m. of 13 batches of islets.

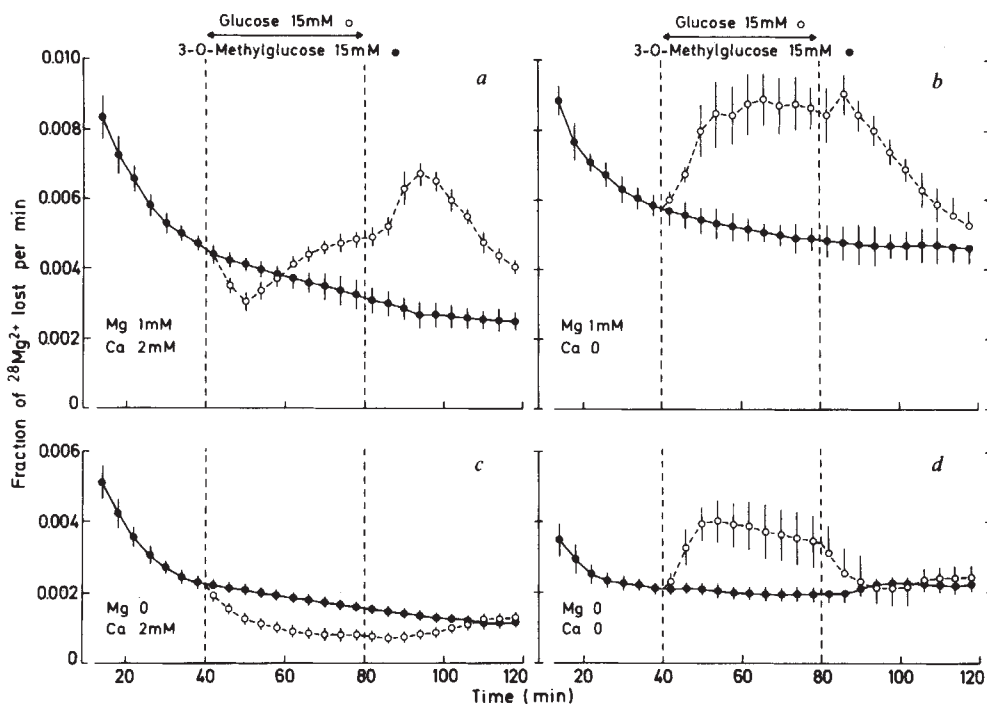
* $P < 0.05$ and $^{\dagger} P < 0.001$ compared with controls; *t*-test for unpaired data.

per islet). It is thus clear that the Mg^{2+} accumulated in response to glucose is localized intracellularly.

All these findings seem to contradict the report⁶ that glucose does not affect total islet content of Mg^{2+} as measured by atomic absorption spectroscopy. However, as our measurements were not made at isotopic equilibrium, they do not provide information on total Mg^{2+} , but simply indicate that glucose accelerates Mg^{2+} turnover in islet cells. This was investigated further by continuously monitoring ^{28}Mg efflux from preloaded islets placed in a perfusion system^{7,8}.

Before addition of the test substances, the fractional rate of ^{28}Mg efflux was higher in the presence (Fig. 1a, b) than in the absence (Fig. 1c, d) of extracellular Mg^{2+} . Whatever the ionic composition of the medium, no significant change in ^{28}Mg efflux occurred on addition or withdrawal of 3-*O*-methylglucose. On the other hand, ^{28}Mg efflux was markedly modified by glucose, depending on the presence or absence of Mg^{2+} and Ca^{2+} in the perfusion medium. Addition of glucose to a control medium containing both cations caused an initial decrease in ^{28}Mg efflux rate and a secondary rise above control values (Fig. 1a). Subsequent removal of the sugar was followed by a further transient increase. When the medium contained Mg^{2+} but no Ca^{2+} , the acceleration of ^{28}Mg efflux triggered by glucose was greater and not delayed (Fig. 1b). These different kinetics are consistent with the greater stimulation of ^{28}Mg influx in the absence of Ca^{2+} . It thus seems reasonable to attribute the increase in efflux rate to a displacement of ^{28}Mg by incoming nonradioactive Mg^{2+} . In the presence of Ca^{2+} and absence of Mg^{2+} , glucose steadily but reversibly decreased the efflux rate of ^{28}Mg (Fig. 1c); this would be compatible with an effect of the sugar on either the Mg^{2+} extrusion mechanisms or the concentration of free cytoplasmic Mg^{2+} . Finally, when both Ca^{2+} and Mg^{2+} were omitted from the extracellular medium, glucose unexpectedly

Fig. 1 Effect of D-glucose (○) and 3-O-methyl-D-glucose (●) on Mg^{2+} efflux from perfused rat islets. The experiments were started in the absence of test substances, which were added, at a concentration of 15 mM, between 40 and 80 min. The concentrations of Mg^{2+} and Ca^{2+} in the perfusion medium are indicated. Values are mean ± 1 s.d. for three experiments with 3-O-methylglucose and four experiments with glucose. The method used has been described previously^{7,8}. Groups of 60–90 islets were first incubated for 120 min in 0.4 ml medium containing 15 mM glucose and 1 mM $^{28}MgCl_2$ (20–50 mCi mmol⁻¹). After three washings with nonradioactive medium, islets were placed in perfusion chambers closed by PVC filters (Polyvic; Millipore) that do not bind ^{28}Mg . Radioactivity in effluent fractions (4 min) and in islets at the end of the experiments was counted by liquid scintillation. After correction for decay, fractional efflux of ^{28}Mg (^{28}Mg released during the time interval/ ^{28}Mg remaining in the tissue during that interval) was calculated for each fraction.



increased ^{28}Mg efflux (Fig. 1d). Attempts to repeat this latter experiment in the presence of EDTA were unsuccessful because of a much higher and variable rate of efflux, possibly due to instability of the plasma membrane.

The experiments described above demonstrate that glucose, the major physiological regulator of the functioning of endocrine pancreas, markedly influences Mg^{2+} fluxes in islet cells. Two lines of evidence suggest that these effects may represent an integral part of the stimulus-secretion coupling. First, we can exclude the possibility that the changes in Mg^{2+} fluxes are merely due to insulin release, as they were not suppressed, but sometimes enhanced, when insulin release was prevented by omission of calcium⁹. Second, none of the effects of glucose could be mimicked by 3-O-methylglucose, which is transported in islet cells like glucose¹⁰, but is not metabolized¹¹, does not depolarize B cells¹² and does not stimulate insulin release¹¹. We are aware of only one other tissue in which a 'physiological' stimulus unequivocally affects ^{28}Mg fluxes. In S49 lymphoma cells, β -adrenergic stimulation inhibits ^{28}Mg uptake without affecting ^{45}Ca uptake¹³. A similar inhibition is also produced by prostaglandin E₁.

It is generally assumed that measurements of ionic fluxes in whole islets are representative of fluxes occurring in the 70–75% of B cells. The antagonistic effects of Ca^{2+} and Mg^{2+} have long been established¹⁴ for insulin release, but are less clear for glucagon release¹⁵. Nevertheless, glucose might alter the movements of the two divalent cations in opposite directions. On the contrary, there exist several analogies, at least qualitative, between the changes in ^{45}Ca fluxes (see ref. 16 for a review) and ^{28}Mg fluxes triggered by glucose: the sugar stimulates influx and net uptake of both cations; it increases the rate of their efflux when they are present in the medium and decreases it when they are absent from the medium (except when Mg^{2+} and Ca^{2+} are omitted simultaneously). In nerve fibres, electrical stimulation also increases Mg^{2+} influx as well as Ca^{2+} influx^{17,18}. These analogies suggest that Mg^{2+} and Ca^{2+} fluxes proceed through common pathways eventually controlled by the membrane potential. Alternatively, the changes in Mg^{2+} fluxes could be secondary to alterations in free cellular Mg^{2+} , brought about by alterations in the cytoplasmic concentration of possible Mg^{2+} chelators. In islet cells stimulated by glucose, the concentration of phosphorylated intermediates of glycolysis increases¹⁹, the total concentration of adenine nucleotides does not change²⁰,

but the proportion of ATP is slightly higher^{19–21}. On the other hand, the concentration of inorganic phosphate decreases substantially, due to an efflux of the anion²¹. It would thus be difficult to predict how these changes will eventually affect Mg^{2+} binding. Although the possible role of variations in the Mg^{2+} buffering capacity of the cytoplasm cannot be excluded, it is unlikely that they are the main cause of the changes in Mg^{2+} fluxes. Thus, the effect of glucose on these metabolic events is not dependent on extracellular Ca^{2+} (refs 19, 21), the omission of which clearly modified the effect of the sugar on Mg^{2+} fluxes.

Further studies are needed to clarify how glucose modifies Mg^{2+} fluxes in islet cells, to establish whether these flux changes are accompanied by changes in free cellular Mg^{2+} and to elucidate their importance for the stimulus-secretion coupling in the various cell types of the endocrine pancreas. It will also be important to determine whether a similar remodelling of Mg^{2+} fluxes occurs in other secretory tissues in which, unlike B cells, the stimulus recognition does not require its metabolic degradation and does not lead to depolarization of the membrane with appearance of electrical activity.

We thank R. Magoche and F. Mathot for assistance. J.C.H. and M.C. are 'Chercheurs Qualifiés' of the FNRS, Brussels. This work was supported in part by grant 3.4552.81 from the FRSM, Brussels.

Received 25 August; accepted 4 November 1982.

- Wacker, W. E. C. *Ann. N.Y. Acad. Sci.* **162**, 717–726 (1969).
- Baker, P. F. & Knight, D. E. *Nature* **276**, 620–622 (1978).
- Baker, P. F. & Knight, D. E. *J. Physiol., Paris* **76**, 497–504 (1980).
- Pace, C. S., Tarvin, J. T., Neighbors, A. S., Pirkle, J. A. & Greider, M. H. *Diabetes* **29**, 911–918 (1980).
- Probst, H. J., Qaim, S. M. & Weinreich, R. *Int. J. appl. Radiat. Isotopes* **27**, 431–441 (1976).
- Andersson, T., Berggren, P. O., Gylfe, E. & Hellman, B. *Acta physiol. scand.* **114**, 235–241 (1982).
- Henquin, J. C. *Nature* **280**, 66–68 (1979).
- Henquin, J. C. *Nature* **271**, 271–273 (1978).
- Grodsky, G. M. & Bennett, L. L. *Diabetes* **15**, 910–913 (1966).
- Hellman, B., Sehlin, J. & Täljedal, L.-B. *Pflügers Arch. ges. Physiol.* **340**, 51–58 (1973).
- Ashcroft, S. J. H. *Ciba Fdn Symp.* **41**, 117–139 (1976).
- Dean, P. M., Matthews, E. K. & Sakamoto, Y. *J. Physiol., London* **246**, 459–478 (1975).
- Maguire, M. E. & Erdos, J. J. *J. biol. Chem.* **255**, 1030–1035 (1980).
- Milner, R. D. G. & Hales, C. N. *Diabetologia* **3**, 47–48 (1967).
- Leclercq-Meyer, V., Marchand, J. & Malaisse, W. J. *Endocrinology* **93**, 1360–1370 (1973).
- Wellheim, C. B. & Sharp, G. W. G. *Physiol. Rev.* **61**, 914–973 (1981).
- Baker, P. F. & Crawford, A. C. *J. Physiol., Lond.* **227**, 855–874 (1972).
- Rojas, E. & Taylor, R. E. *J. Physiol., Lond.* **252**, 1–27 (1975).
- Idahl, L. A. *Hormone Metab. Res. Suppl.* **10**, 20–26 (1980).
- Malaisse, W. J. *et al. Diabetologia* **16**, 331–341 (1979).
- Bukowiecki, L., Trus, M., Matschinsky, F. M. & Freinkel, N. *Biochim. biophys. Acta* **583**, 370–377 (1979).
- Henquin, J. C. & Lambert, A. E. *Am. J. Physiol.* **228**, 1669–1677 (1975).

Tracing the use of torture: electrically induced calcification of collagen in pig skin

T. Karlsmark*, L. Danielsen*†, H. K. Thomsen†, O. Aalund‡, O. Nielsen§, K. G. Nielsen§, E. Johnson|| & I. K. Genefke§

* Dermatology Clinic, Municipal Hospital, Copenhagen, Denmark

† Department of Pathology, Hvidovre Hospital, Copenhagen, Denmark

‡ Royal Veterinary and Agricultural University, Copenhagen, Denmark

§ Anti-Torture Research Index, Trondhjems gate 11, Copenhagen, Denmark

|| Physics Laboratory II, H. C. Ørsted Institute, University of Copenhagen, Denmark

Reports¹⁻³ of the use of electrical torture are usually denied, so that there is an urgent need of diagnostic methods to distinguish the consequences of electrical torture from other superficial injuries. The 'electrical group' of Anti-Torture Research (ATR) has accordingly studied the possibility of distinguishing between the sequelae of electrical and heat injury. Among the reported differences⁴⁻⁶ of the two types of injuries is the occurrence in epidermis, vessel walls and sweat glands of vesicular nuclei exclusively in electrically injured skin. On the basis of experiments with anaesthetized pigs, we now report that the late sequelae of electrical injury appear to include the deposition of calcium salts beneath the area of an electrical cathode.

Seven fully anaesthetized pigs (12–35 kg) were used in the experiments. One pig was exposed in the first experiment, and two pigs in each of the following three experiments. Each pig was exposed to heat energy on the left side of the body and to electrical energy on the right side⁷. The areas of exposure on the left side were positioned symmetrically to those on the right side. Except for the first experiment, where 50 Hz a.c. was not used, all electrical experiments included d.c. and 50 Hz a.c. Both types of energy were applied in consecutive exposures during one period of anaesthesia.

† To whom correspondence should be addressed.

Skin biopsies were taken from the centre of lesions 6–65 days after the exposure and examined by light microscopy following staining with haematoxylin/eosin and von Gieson-Hansen/Alcian blue or toluidine blue staining (see Fig. 1 legend for experimental details). These stainings have previously been used as routine methods, because they have proved effective in revealing abnormal vesicular nuclei in electrically injured skin⁴⁻⁶ and to visualize dermal fibres. As basophilic material was found in the dermis of the haematoxylin/eosin-stained sections of biopsies obtained from the cathode areas 6 days after exposure, deposition of calcium salts was anticipated. Calcium appears bluish-black in this staining. However, to overcome the lack of specificity and sensitivity of that reaction the presence of calcium was confirmed by using Alizarin red S, which has been reported to be specific and sensitive for calcium⁸. Alizarin red S was then regularly used in the subsequent series. The occurrence of calcium deposits was registered together with the structures involved in the calcification.

To confirm the presence of calcium salts further and to reveal their exact localization, one 3-mm punch biopsy was taken for ultrastructural examination from the periphery of the exposure zone of one cathode area of pig no. 2 15 days after exposure to 78 J of d.c. Biopsy specimens were embedded in epoxy resin and sections stained with Alizarin red S. Sections showing positive calcium reaction were selected for X-ray fluorescence microanalysis (Fig. 2), electron diffraction analysis (Fig. 3) and transmission electron microscopy (Fig. 1c, d). Ultrathin sections were cut with an LKB ultramicrotome. Sections for transmission electron microscopy were stained using a combined technique with uranyl acetate and lead citrate^{9,10}, while sections for electron diffraction analysis and microanalysis were left unstained. This preparation technique has been shown to induce only insignificant alterations of crystalline calcified deposits¹¹.

In scars following heat damage we found deposits of calcium salts in hair follicles (Table 1a) and in structures of the size and shape of fat cells usually appearing in clusters (Table 1c). However, it was not possible to evaluate whether structures other than fat cells (for example, sweat glands, sebaceous glands or vessels) were covered by the calcium deposits (Table 1a, b). The biopsy listed as a control biopsy and containing calcium deposits (Table 1a) might have been misplaced and represent a heat-damaged area as it contained young connective tissue located as in the heat biopsies.

An accumulation of calcium salts on dermal collagen fibres was observed in scars following electrical injury (Tables 1, 2).

Table 1 Calcium deposits in electrically and heat-influenced dermis from pigs at various times after injury, Alizarin red S staining

Total no. of biopsies examined	Influence	Energy range (mean) (J)	No. of biopsies with calcium deposits				Signs of rejection
			Total	Collagen fibres	Hair follicles	Unclassified structures*	
a, 15 Days after injury							
4	Anode	78-91 (82)	1	1	1	1	0
4	Cathode	78-91 (82)	4	2	1	3	0
10†	a.c.	200-285 (241)	3	0	2	2	0
16	Unexposed	—	1	0	0	1	0
16	Heat	43-86 (61)	7	0	4	4	0
b, 21 Days after injury							
4	Anode	78-96 (83)	0	0	0	0	0
4	Cathode	78-96 (83)	4	3	0	4	2
8	a.c.	200-225 (206)	0	0	0	0	0
16	Unexposed	—	0	0	0	0	0
16	Heat	49-100 (69)	11	0	0	11	0
c, 2 Months after injury							
4	Anode	79-85 (81)	0	0	0	0	0
4	Cathode	79-85 (81)	3	3	0	0	0
8	a.c.	200 (200)	0	0	0	0	0
16	Unexposed	—	0	0	0	0	0
16	Heat	42-115 (67)	6	0	1	5	0

* Structures which could not be identified with certainty, some of which have been evaluated as sweat glands and ducts, sebaceous glands or vessels.

† Including two biopsies taken in two areas of pig no. 3 from a region showing macroscopic alterations between the two exposure zones.

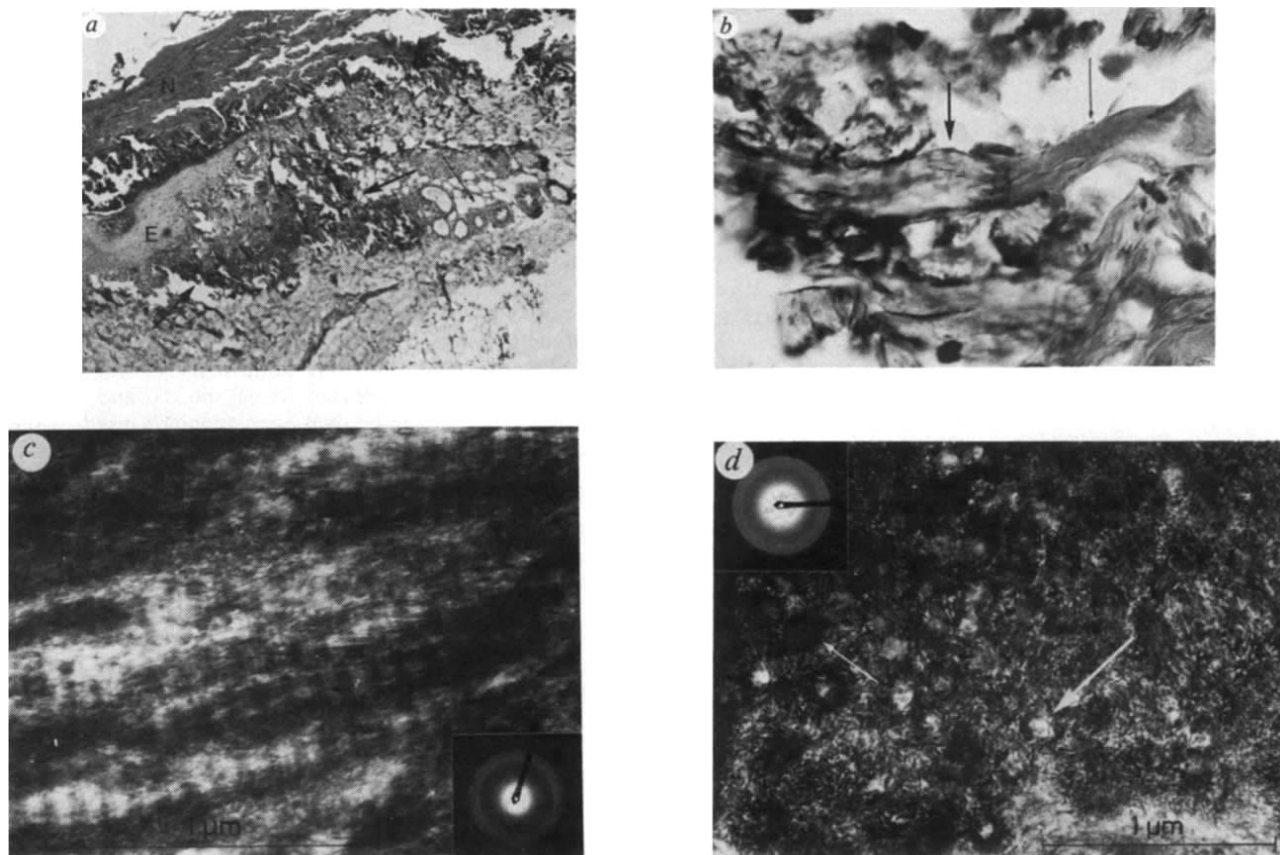


Fig. 1 Effects of heat and electrical injury on skin histology and ultrastructure in anaesthetized pigs. Each pig underwent heat exposure on the left side of the body and electrical energy exposure on the right side. The energy was supplied on each side to four areas of abdominal skin and one area of back skin. The transfer of energy to each area was done via two circular and slightly convex metal bodies with a diameter of 12 mm and a centre distance of 50 mm mounted on a plate of insulating material⁴. During heat exposure the metal bodies were thermostatically regulated to 55–90 °C (ref. 7). The amount of energy applied was 42–148 J in heat experiments and 77–96 J and 200–285 J in experiments with d.c. and a.c., respectively. Higher energy doses of a.c. than d.c. were chosen because preliminary experiments had indicated that a.c. is required at higher energy doses than d.c. to produce alterations specific for electrical damage (vesicular nuclei). To ensure good contact during the experiments the electrodes were fixed to the skin by slight manual pressure. The exposure time was 40 s in all experiments. For light microscopy a 5-mm punch biopsy was taken from the centre of the lesions, 6–65 days after the exposure, each pig being once more fully anaesthetized. Several biopsies from unexposed abdominal skin areas of each pig were also included in the material. The biopsies from the back skin were kept for later use, while all the other biopsies were given random numbers, cut in serial sections and evaluated in blind studies, the biopsies from one pig being examined at a time by means of Alizarin red S, haematoxylin/eosin and von Gieson-Hansen/Alcian blue or toluidine blue staining. For transmission electron microscopy and diffraction analysis one 3-mm punch biopsy was used (see Figs 2 and 3 legends). *a*, Cathode area 6 days after exposure to 78 J d.c. A band of dark collagenous tissue (thick arrows) is seen extending from the lower side of the newly formed epidermis (E) down to a group of sweat glands (thin arrow). A layer of necrotic tissue (N) covers the epidermis. Haematoxylin/eosin staining. $\times 30$. *b*, Cathode area 15 days after exposure to 90 J by d.c. Collagen fibre showing a pale centre mixed with and framed by dark narrow lines to the left (thick arrow) and normal stain to the right (thin arrow). Haematoxylin/eosin staining. $\times 130$. *c*, Cathode area 15 days after exposure to 78 J by d.c. Dense accumulations on longitudinal sectioned collagen fibrils. The inset shows the selected area diffraction pattern of the deposit indicating that the deposit is of apatite type¹². Uranyl acetate/lead citrate staining. $\times 47,000$. *d*, Cathode area 15 days after exposure to 78 J by d.c. Cross-section of calcified collagen fibrils shows apatite crystals around (thick arrow) as well as inside (thin arrow) collagen fibrils. Uranyl acetate/lead citrate staining. $\times 30,000$.

Table 2 Calcium deposits in electrically and heat-influenced dermis from pig no. 1 6 days after the injury, haematoxylin/eosin staining

Total no. of biopsies examined	Influence	Energy range (mean) (J)	No. of biopsies with calcium deposits				Signs of rejection
			Total	Collagen fibres	Hair follicles	Unclassified structures	
4	Anode	77-81 (79)	0	0	0	0	0
5*	Cathode	77-81 (79)	4*	4*	0	0	0
8	Unexposed	—	0	0	0	0	0
8	Heat	48-148 (89)	0	0	0	0	0

* Including one biopsy from the back skin area stained with Alizarin red S and not examined in 'blind' studies.

This alteration was associated with the cathode area. The calcified collagen fibres observed in a single anode area (Table 1a) might represent a different structure, such as keratin fibres of a hair follicle. The collagen calcification occurred in bands apparently following the sweat ducts (Fig. 1a). Although on day 6 after injury these collagen fibres were dark blue in

haematoxylin/eosin stainings, they had lost their stainability 9 days later. Alizarin red S staining was positive on both occasions. Some of the collagen fibres showed a distinct boundary zone between unaffected and calcified areas of the fibres (Fig. 1b), suggesting that accumulation had occurred on preexisting mature fibres. Fifteen days after exposure the

calcified areas were split up into numerous small oblong units having the diameter of collagen fibres. Occasionally, the calcified units were difficult to recognize as collagen-related structures. Such sites are indicated as unclassified structures (Table 1a, b). A cellular inflammatory reaction was seen around the calcified fibres. In these areas the epidermis was acanthotic, with elongation of the rete ridges often surrounding calcified material. Such areas were initially evaluated as representing calcified sweat ducts and are indicated as unclassified structures in Table 1a, but should have been indicated as signs of rejection. Three weeks after the exposure the rete ridges were seen extending to the level of the subcutaneous tissue and signs of rejection of calcified material became evident (Table 1b). Two months after the exposure no signs of rejection were observed, but numerous calcified units as described above were still present in the dermis (Table 1c). This type of calcification was never observed in heat lesions, but a few electrically influenced biopsies, mainly those injured by a.c., contained calcium deposits as seen in heat-damaged areas (Table 1a), probably caused by the concomitant development of heat during the passage of the electrical current.

Ultrastructurally, electron-dense accumulations were observed around and within collagen fibrils in the cathode area (Fig. 1c, d), and contained granules and thin needles of high electron density. X-ray microanalysis showed the presence of calcium and phosphorus (Fig. 2) inside the accumulations, while these elements were absent from the control spectra. The presence of sharp rings in the diffraction patterns from the accumulations of both stained and unstained sections (Fig. 3) provided

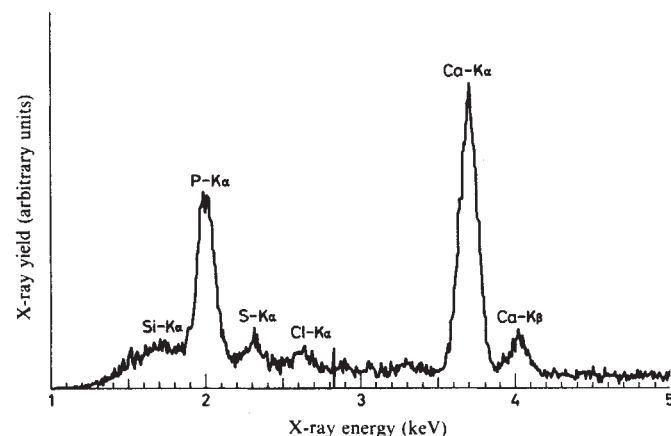


Fig. 2 Electron-induced X-ray fluorescence spectrum from a dense calcified deposit obtained from a punch biopsy from the periphery of the exposure zone of a cathode area in a pig 15 days after electrical injury. The biopsy was fixed at 4 °C overnight in 6% glutaraldehyde in Veronal (0.03 M) acetate (0.03 M) buffer pH 7.4 containing 7.5% sucrose, minced and post-fixed in 1% osmic acid for 1 h. After stepwise dehydration in increasing concentrations of ethanol biopsy specimens were embedded in epoxy resin. Sections (1 µm thick) were stained with Alizarin red S; sections from two blocks showed positive calcium reaction in parts of the dermis, and these areas were selected. Transmission electron microscopy and diffraction analysis (see Figs 1c, d and 3) were carried out using a JEOL 100U transmission electron microscope at 100 kV acceleration voltage. The elemental composition of the calcified deposits was determined in a Cambridge 180 scanning electron microscope equipped with a modified transmission stage and using a Si(Li) energy discriminating detector (Kevex 163). A beryllium window covering the detector allows detection only of elements having atomic numbers higher than that of sodium. X-ray spectra were recorded from unstained sections mounted on copper grids coated with Formvar and carbon. Control spectra were obtained from areas of the sections containing embedding material only. The microscope was operated at 20 kV acceleration voltage and the probed areas were ~4 µm². Prominent calcium and phosphorus peaks are seen. The chlorine peak is due to tissue background, and the silicon and sulphur peaks are due to instrumental background.

evidence of crystallinity. The crystal lattice spacings obtained from the observable diffraction rings in Fig. 3 are listed in Table 3. Combined with the elemental analysis (Fig. 2), the diffraction analysis^{12,13} unambiguously shows that the accumulations consist of crystalline calcium apatite. A final distinction between calcium hydroxyapatite and calcium fluoroapatite cannot be made, either from X-ray microanalysis which cannot detect fluorine, or from diffraction analysis, as the two materials have virtually identical structures. This can be seen from Table 3 where lattice spacings and equivalent Miller indices for the two materials (as reported in the American Society for Testing and Materials (ASTM)¹⁴) are included.

Thus, calcification of preexisting dermal collagen fibres was found a few days after injury by d.c. Although the amount of deposited calcium was reduced with time, considerable calcium was still present 2 months after the injury. Bone formation was not observed. Skin damaged by a.c. or heat energy contained calcium deposits which were probably related to cellular structures. As the production of characteristic morphological changes in electrically damaged skin requires higher doses of a.c. than of d.c., a higher dose of a.c. than applied here may produce collagen calcification.

Deposition of calcium salts in high concentration on the skin surface results in calcification of dermis when the epidermal integrity is disturbed^{15,16}. The calcium salts tend to precipitate on collagen fibres and histologically may be indistinguishable from electrically induced calcification. Ultrastructural investigations of such calcium deposits have not been performed.

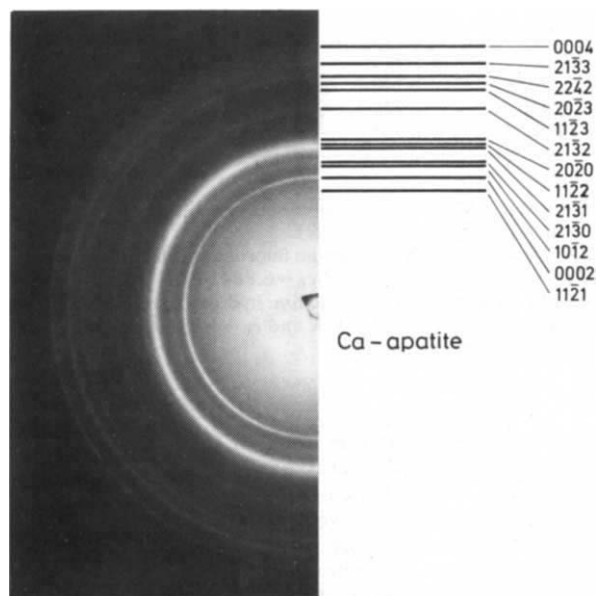


Fig. 3 Selected area diffraction pattern from calcified deposit in unstained ultrathin section. Crystallinity and crystal structure of calcified accumulations were unambiguously assessed from diffraction analysis of the selected areas^{12,13}. Use of unstained specimens prevents a high background level in the diffraction patterns from the staining materials. Good contrast and visibility of the mineral accumulations are ensured from their relatively high electron scattering power, and identification of tissue structures could be reasonably well achieved due to weak tissue staining from the osmium used in post-fixation. The diffraction constant of the electron microscope was determined from ring diameters in diffraction patterns of evaporated, thin polycrystalline aluminium films. By recording sets of diffraction patterns from calcified deposits and from the aluminium film with unchanged lens excitations between exposures, the planar spacings were determined with a standard deviation not exceeding 1%¹³. The rings are evidence of crystallinity and the diffraction analysis shows that the material is calcium apatite. Individual rings are labelled using the four-index Miller-Bravais notation.

Table 3 Experimental electron diffraction data presenting planar distances $d(hkil)$ obtained from a calcified deposit in an unstained ultrathin section, compared with X-ray diffraction data for calcium fluoroapatite and calcium hydroxyapatite¹⁴

Miller indices* $hkil$	ASTM data Fluoroapatite† $d(hkil)$ Å	Hydroxyapatite‡ $d(hkil)$ Å	Experimental data $d(hkil)$ Å
10 $\bar{1}$ 0	8.12	8.17	
10 $\bar{1}$ 1	5.25	5.26	
11 $\bar{2}$ 0	4.684	4.72	
20 $\bar{2}$ 0	4.055	4.07	
11 $\bar{2}$ 1	3.872	3.88	3.87§
20 $\bar{2}$ 1	3.494	3.51	
0002	3.442	3.44	3.44
10 $\bar{1}$ 2	3.167	3.17	3.18
21 $\bar{3}$ 0	3.067	3.08	3.05
21 $\bar{3}$ 1	2.800	2.814	2.84
11 $\bar{2}$ 2	2.772	2.778	2.79§
30 $\bar{3}$ 0	2.702	2.720	
20 $\bar{2}$ 2	2.624	2.631	2.63
30 $\bar{3}$ 1	2.517	2.528	
21 $\bar{3}$ 2	2.289	2.296	2.29
31 $\bar{4}$ 0	2.250	2.262	
22 $\bar{4}$ 1	2.218	2.228	
31 $\bar{4}$ 1	2.140	2.148	
30 $\bar{3}$ 2	2.128	2.134	
11 $\bar{2}$ 3	2.061	2.065	2.07§
40 $\bar{4}$ 0	2.028	2.040	
20 $\bar{2}$ 3	1.997	2.000	2.00§
22 $\bar{4}$ 2	1.937	1.943	1.94
31 $\bar{4}$ 2	1.884	1.890	
32 $\bar{5}$ 0	1.862	1.871	
21 $\bar{3}$ 3	1.837	1.841	1.84
32 $\bar{5}$ 1	1.797	1.806	
41 $\bar{5}$ 0	1.771	1.780	
40 $\bar{4}$ 2, 30 $\bar{3}$ 3	1.748	1.754	1.72
0004, 41 $\bar{5}$ 1	1.722	1.722	
10 $\bar{1}$ 4	1.684	1.684	

* The four-index notation, commonly used for hexagonal crystals. The index i is given by $i = -(h+k)$.

† ASTM card no. 15-876. Calcium fluoroapatite has hexagonal crystal structure with $a_0 = 9.3684$ Å and $c_0 = 6.8841$ Å.

‡ ASTM card no. 9-432. Calcium hydroxyapatite has hexagonal crystal structure with $a_0 = 9.418$ Å and $c_0 = 6.884$ Å.

§ These rings are very weak.

Calcium deposits occur in areas of scar formation but are usually not present until several years after the injury¹⁷. Usually, but not always, there is concomitant formation of bone tissue¹⁸, which has been observed 3 weeks after the injury (see ref. 18). In cases without bone tissue the calcium deposits are related to hyalinized dense connective tissue and tend to increase in subsequent years¹⁷. Calcium deposits in combination with local conditions of the skin other than scar tissue have been reported to be associated with cellular structures^{19,20}, and sometimes to contain bone tissue¹⁸.

Calcium deposits in the skin unrelated to any other conditions (idiopathic calcinosis cutis)^{18,21} are not entirely restricted to the collagen fibres, and ultrastructurally they differ from those of electrically injured skin. Calcium deposits in the skin related to systemic diseases²²⁻²⁵ show several morphological patterns, but again they differ from those reported here for electrically induced lesions. Thus, the collagen calcification described here is unique and in the absence of a history of skin exposure to calcium salts, highly indicative of electrical injury. Ultrastructural demonstration of intrafibrillar small needles constitutes further evidence of such injury.

We thank Hans Jørgen Hansen for contributions to the X-ray microanalysis and Birte Christiansen, Mirosława Gonik, Elin Just Olesen, Inger Jensen and Poul Henriksen for technical assistance. This work was supported by the Danish MRC.

Received 6 September; accepted 15 September 1982.

1. Amnesty International Report on Torture (Duckworth, London, 1975).
2. Kjaersgaard, Aa. R. & Genefke, I. K. *Ugeskr. Laeg.* **139**, 1057-1059 (1977); in *Evidence of Torture* (Amnesty International Publications, London, 1977).
3. Cathcart, L. M., Berger, P. & Knazan, B. *Can. med. Ass. J.* **121**, 179-184 (1979).
4. Danielsen, L. *et al. Forensic Sci. Int.* **12**, 211-225 (1978).
5. Aalund, O. *Acta med. legalis socialis* **30**, 33-41 (1980).
6. Thomsen, H. K. *et al. Forensic Sci. Int.* **17**, 133-143 (1981).
7. Nielsen, K. G., Nielsen, O. & Thomsen, H. K. *Forensic Sci. Int.* **17**, 203-209 (1981).
8. Dahl, L. K. *Proc. Soc. exp. Biol. Med.* **80**, 474-479 (1952).
9. Reynolds, E. S. *J. Cell. Biol.* **17**, 208-212 (1963).
10. Watson, M. L. *J. Biophys. Biochem. Cytol.* **4**, 475-485 (1958).
11. Johnson, E. & Otkjaer-Nielsen, A. *Proc. 7th Eur. Congr. Electron Microscopy, The Hague, 1980* Vol. 2 (eds Brederoc, P. & de Priester, W.) 416-417 (1980).
12. Otkjaer-Nielsen, A., Johnson, E., Hentzer, B., Danielsen, L. & Carlsen, F. *J. invest. Derm.* **69**, 376-378 (1977).
13. Hentzer, B., Nielsen, A. O. & Johnson, E. *Br. J. Derm.* **102**, 675-685 (1980).
14. *ASTM Inorganic Index to the Powder Diffraction File*, Powder Diffraction Index Set 17i (ASTM, Philadelphia, 1967).
15. Clendenning, W. E. & Auerbach, R. *Archs. Derm.* **89**, 360-363 (1964).
16. Sneddon, I. B. & Archibald, R. *Mel. Br. J. Derm.* **70**, 211-214 (1958).
17. Hogan, V. M. & Conway, H. *Plastic and Reconstructive Surg.* **33**, 559-563 (1964).
18. Roth, C. S. I., Stowell, R. E. & Helwig, E. B. *Archs Path.* **76**, 44-54 (1963).
19. Woods, B. & Kellaway, T. D. *Br. J. Derm.* **75**, 1-11 (1963).
20. Hashimoto, K., Nelson, R. G. & Lever, W. F. *J. invest. Derm.* **46**, 391-408 (1966).
21. Cornelius, C. E., Tenenhouse, A. & Weber, J. C. *Archs. Derm.* **98**, 219-229 (1968).
22. Danielsen, L. *Acta derm. Suppl.* **83**, 1-79 (1979).
23. Greenebaum, E. *Hum. Path.* **11**, 287-289 (1980).
24. Winkelmann, R. K. & Keating, F. R. Jr *Br. J. Derm.* **83**, 263-268 (1970).
25. Putkonen, T. & Wangel, G. A. *Dermatologica* **118**, 127-144 (1959).

Magnetic bones in human sinuses

R. Robin Baker, Janice G. Mather & John H. Kennaugh

Department of Zoology, University of Manchester, Manchester M13 9PL, UK

Studies on the interaction of magnetic fields and biological organisms¹ have centred on the influence of applied magnetic fields on the physiology and behaviour of organisms, including humans, and a search for magnetic sources within the organisms themselves. Evidence continues to accumulate that a wide range of organisms, from bacteria to vertebrates, can detect and orient to ambient magnetic fields (for examples see refs 2-4). Since the discovery that magnetic orientation by bacteria was due to the presence within the organism of magnetic particles of the ferric/ferrous oxide, magnetite^{5,6}, the search has begun for other biogenic deposits of inorganic magnetic material and ways in which the possession of such material might confer on the organism the ability to orient to ambient magnetic fields. Such magnetic material, often identified as magnetite, has been discovered in bees⁷, homing pigeons^{8,9}, dolphins¹⁰ and various other organisms⁴, including man¹¹. A variety of hypotheses for the use of magnetite in magnetic field detection have been proposed^{12,13}. We report here that bones from the region of the sphenoid/ethmoid sinus complex of humans are magnetic and contain deposits of ferric iron. The possible derivations and functions of these deposits are discussed.

The search for magnetic material in the tissue of animals has now become relatively commonplace⁴, and standard protocols have been developed¹³. Basically, the method involves testing tissue samples for magnetic remanence in a magnetometer. If the tissue is exposed to a sufficiently strong magnetic field immediately before testing, any permanently magnetic particles within the sample take on common alignment and the measured induced magnetic remanence (IRM) reflects the total amount of permanently magnetic material present. Some workers¹⁴ have also begun a histological search for ferric iron using the Perl reaction¹⁵. We have used both approaches in the present study.

Tissue samples were obtained from recently dead (1-2 days) U22 subjects using a chromium-plated, stainless steel hammer saw. Samples were placed in 5% neutral formalin and tested for magnetic remanence within 5 days using a par-astatic

magnetometer. A minute or so before testing, the tissue was exposed to a strong electromagnetic field of $\sim 2,000$ G.

For histological examination, after standard wax embedding, sectioning was carried out on a MSE sledge microtome, and sections removed using pressure-sensitive tape. Sections were placed on slides coated with Mayer's albumin and dried in an oven at 50°C for 2 days. Tape and adhesive were removed by immersing slides first in acetone and then chloroform before taking the sections through xylene and down through the alcohol series to distilled water. They were then stained for ferric iron using the Perl reaction. This involved placing the sections in a warm solution of equal parts of 4% potassium ferrocyanide and 4% HCl for no more than 10 min in an oven at $\sim 60^\circ\text{C}$ (ref. 15), though the same results were obtained by leaving the slides in the solution for 30 min at room temperature. Slides were washed in distilled water, dehydrated in the alcohol series, lightly counterstained in eosin, cleared in xylene, and then mounted in Depex. Any ferric iron showed as deep blue coloration. Control tests showed that magnetite particles from geological sources stained deep blue after passing through this sequence.

Magnetometric measurements on 'soft' tissues from the human head revealed no areas of significant magnetic remanence (that is, no readings greater than twice the background noise level where background includes the IRM of the plastic container holding the tissue). Specimens of skull, rib and sphenoid bone also gave readings less than twice background levels. In the thin but very hard bones that form the walls of the sphenoid/ethmoid sinus complex, however, magnetic remanence was obtained that was consistently greater than twice background (range 2 to 13 times background; median 9). The range of IRM (e.m.u. g^{-1}) for all tissues tested is shown in Table 1.

Histological examination was carried out on tissue from all of the regions tested magnetometrically except for the bone of the skull and the sphenoid itself, which we were unable to section. Although a scattering of iron-staining material was found throughout the soft tissues and in bone marrow, extensive concentrations of ferric iron were found only within the bone of the sphenoid/ethmoid sinus complex (Fig. 1). These concentrations were in the form of a continuous layer of iron-staining material, apparently particulate, approximately $2\ \mu\text{m}$ thick and about $5\ \mu\text{m}$ beneath the surface of the bone. The bone itself is about $200\ \mu\text{m}$ thick. Although it is difficult to follow the orientation of our material through from intact sinus to mounted section, our impression is that the iron layer is most conspicuous beneath the surface of the bone that forms the inner wall of the sinus. Occasionally, however, there is an iron layer about $5\ \mu\text{m}$ beneath both inner and outer surfaces.

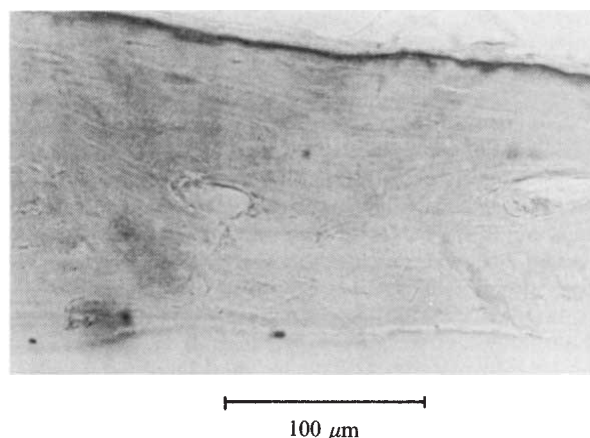


Fig. 1 Section through thin bone forming the wall of the sphenoid sinus of a 57-yr-old woman showing a continuous layer of a ferric iron material $\sim 5\ \mu\text{m}$ below the surface. Stained using the Perl reaction.

We have examined the sinus bones of five subjects and in four have found IRM greater than two times the background level. The single exception (omitted from Table 1) was suffering from anaemia on death and histological examination of the sinus bones revealed no iron.

The possibility exists that our magnetometric and histological results are artefacts due to contamination either during each subject's lifetime or during experimental processing after death. However, after careful consideration of several possible sources of contamination (for example, nose bleeds, magnetic dust, dissecting and sectioning instruments) we consider it highly unlikely that the magnetic deposits are due to either internal or external contamination.

If we accept that human sinus bones are naturally magnetic, at least three tentative hypotheses for the observed deposits may be offered. One possibility is that the sinus bones are part of a storage or dumping site for iron. The apparent absence of deposits in an anaemic subject suggests that the material is withdrawn if the subject becomes iron-deficient. A second possibility is that the magnetic deposits are concerned with magnetic field detection. Experiments suggest that humans can detect direction by reference to the ambient magnetic field^{16,17} although the conclusion is still a matter of contention¹⁸. The sphenoid/ethmoid sinuses coincide with the region deduced from orientation experiments to be the site of a magnetoreceptor¹⁷. A third possibility is that the magnetic material is involved in some way with growth and repair of the bones. If so, we should anticipate similar deposits in other bones, though not

Table 1 Inducible magnetic remanence in human tissue

	No. of individuals	No. of samples	Median magnetic remanence (range) (e.m.u. $\times 10^{-6}$ per g of tissue)
Soft tissue			
Brain	2	5	-0.004 (-0.007-0.003)
Olfactory lobe	1	1	0.008
Optic chiasma	1	1	0.009
Dura (top of skull)	3	4	0.215 (-1.790-0.900)
Dura (sphenoid)	3	5	1.540 (-0.010-6.520)
Pineal	4	4	2.360 (-0.234-8.380)
Bone			
Rib	1	1	0.780
Skull	1	1	1.150
Nasal conchae	1	1	2.050
Sphenoid	1	1	0.190
Walls of:			
Sphenoid sinus (+ethmoid fragments)	2	4	16.110 (5.810-24.000)
Sphenoid sinus	4	7	13.000 (3.020-31.580)

Magnetic remanence was calculated as $(TC - C)/W$, where TC = e.m.u. of tissue + container, C = e.m.u. of container and W = weight of tissue (g). All readings were taken after exposure to a field of 2,000 G. Overlap in the remanence of pineal and dura tissue with that of bone from the sinus region is more apparent than real and is due to scaling-up within-noise readings for small samples of light tissue and scaling-down above-noise readings for heavy samples of bone.

necessarily with magnetic remanence levels detectable by our magnetometer. Some authors report that the rate of healing of broken bones may be influenced by applied electromagnetic fields¹⁹.

Such interpretations need not be mutually exclusive: iron may be dumped or stored in bones and may impart magnetic properties, aiding growth and repair in the case of breakage, and may be exaggerated in the bones of the sinus region for use in orientation. Whatever the final explanation, the discovery of magnetic material in the bones of the human sinus region seems to open novel avenues for biomagnetic research.

We thank Professor Yates, Dr M. I. Wright and the staff of the Pathology Department of Manchester University for providing us with material and advice, and Dr Peter Dagley and the Geophysics Department, University of Liverpool, for the use of their magnetometer. This work was supported by SERC grant GR/B74337 to R.R.B.

Received 28 June; accepted 11 November 1982.

- Williamson, S. J. & Kaufman, L. J. *Magn. magn. Mater.* **22**, 129–202 (1981).
- Schmidt-Koenig, K. & Keeton, W. T. (eds) *Animal Migration, Navigation and Homing* (Springer, Heidelberg, 1978).
- Mather, J. G. & Baker, R. R. *Nature* **291**, 152–155 (1981).
- EOS* **62**, 849–850 (1981); **63**, 156 (1982).
- Blakemore, R. P. *Science* **190**, 377–379 (1975).
- Frankel, R. B., Blakemore, R. P. & Wolfe, R. S. *Science* **203**, 1355 (1979).
- Gould, J. L., Kirschvink, J. L. & Deffeyes, K. S. *Science* **201**, 1026–1028 (1978).
- Walcott, C., Gould, J. L. & Kirschvink, J. L. *Science* **205**, 1027–1029 (1979).
- Presti, D. & Pettigrew, J. D. *Nature* **285**, 99–101 (1980).
- Zoeger, J., Dunn, J. R. & Fuller, M. *Science* **213**, 892–894 (1981).
- Kirschvink, J. L. *J. exp. Biol.* **92**, 333–335 (1981).
- Yorke, E. D. *EOS* **62**, 849 (1981).
- Kirschvink, J. L. & Gould, J. L. *Biosystems* **13**, 181–201 (1981).
- Walcott, B. & Walcott, C. in *Avian Navigation* (eds Papi, F. & Wallraff, H. G.) 338–343 (Springer, Heidelberg, 1982).
- Hutchison, H. E. *Blood* **8**, 236–240 (1953).
- Baker, R. R. *Science* **210**, 555–557 (1980).
- Baker, R. R. *Human Navigation and the Sixth Sense* (Hodder & Stoughton, London, 1981).
- Gould, J. L. & Able, K. P. *Science* **212**, 1061–1063 (1981).
- Bassett, C. A., Pawluk, R. J. & Pilla, A. A. *Science* **184**, 575–577 (1974).

Thy-1 cDNA sequence suggests a novel regulatory mechanism

Tetsuya Moriuchi, Hsui-Ching Chang, Roger Denome & Jack Silver

Department of Microbiology and Public Health,
Michigan State University, East Lansing, Michigan 48824, USA

Thy-1 was originally defined in mice as a cell-surface alloantigen of thymus and brain with two allelic forms, Thy-1.1 and Thy-1.2 (ref. 1). Subsequently, the Thy-1.1 alloantigenic determinant was identified in rats². In both species, Thy-1 is present in large amounts on thymus and brain cells³ and in smaller quantities on fibroblasts⁴, epidermal cells⁵, mammary glands⁶ and immature skeletal muscle⁷. In many of these tissues the level of Thy-1 expression changes dramatically during cell differentiation. The molecules expressing the Thy-1 antigenic determinant have been isolated from rat and mouse brain cells and have been shown to have a molecular weight of 17,500 (ref. 8). One-third of the Thy-1 molecule is carbohydrate and the remainder is a polypeptide of 111 amino acids whose sequence has been fully determined⁹. We report here the isolation and characterization of a cDNA clone encoding the rat thymus Thy-1 antigen but find that the DNA sequence ends prematurely at a position corresponding to amino acid 103. It appears to be a complete transcript, however, as the last codon is followed directly by a poly(A) tract.

Poly(A) containing RNA was isolated from W/Fu rat thymocytes, and cDNA was synthesized by reverse transcription using oligo(dT)_{12–18} primer. Double-stranded cDNA was synthesized as described previously¹⁰. A cloned cDNA library was constructed by inserting the total cDNA population into

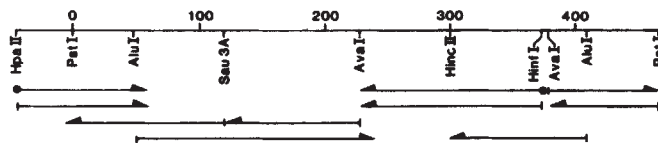


Fig. 1 Partial restriction map and strategy for sequencing Thy-1 cDNA clone, pT64. The entire cDNA clone was cleaved at selected sites with restriction endonucleases and fragments corresponding to the insert were purified by acrylamide gel electrophoresis and electroelution. Fragments were treated with calf intestinal or bacterial alkaline phosphatase and ³²P-labelled at both 5' ends with T4 polynucleotide kinase as described¹³. Alternatively, some fragments were ³²P-labelled at both 3' ends with DNA polymerase²⁶. Double-end-labelled fragments were separated by acrylamide gel electrophoresis, electro-eluted and subjected to partial chemical degradation sequence analysis as described by Maxam and Gilbert¹³. Partially cleaved fragments were separated on 8, 15 and 20% acrylamide gels. The extent of sequence determined from each fragment is indicated by the length of the arrow. The closed circles and vertical lines at one end of each arrow indicate labelling at the 3' end or 5' end, respectively. The sequencing strategy used here allowed us to sequence both DNA strands completely.

the *Pst*I site of pBR322 using poly(dG).poly(dC) homopolymeric extensions¹¹. Colonies containing cDNA were screened¹² with a synthetic ³²P-labelled oligodeoxynucleotide (17-mer) mixture composed of all 32 possible sequence permutations corresponding to amino acids 82–87 of Thy-1. Two hybridization-positive clones were isolated from ~10,000 colonies. One clone (pT64) contained cDNA encoding the Thy-1 antigen and its entire sequence was determined by the procedure of Maxam and Gilbert¹³. The overall structure of the cDNA clone pT64 and the sequencing strategy used are shown in Fig. 1.

The DNA sequence contains a 45-base 5'-untranslated region followed by 57 bases coding for a presumptive leader peptide of 19 amino acids starting at the first methionine codon (numbered –19 in Fig. 2). The leader sequence exhibits several features characteristic of leader peptides present at the amino terminus of membrane bound proteins¹⁴; it contains many hydrophobic amino acids (11 non-polar residues) and terminates in a residue with a small neutral side chain (glycine). The remainder of the predicted amino acid sequence is in agreement with the published amino acid sequence for rat brain Thy-1⁹. There are, however, several unusual features at the 3' end of the Thy-1 cDNA clone. The DNA sequence ends prematurely at a position corresponding to amino acid 103, 8 amino acids earlier than predicted from the protein sequence, but is followed directly by a poly(A) tract. There is, however, no termination codon and, furthermore, a presumptive polyadenylation signal, 5'-AATAAA-3' (ref. 15), which is part of the coding sequence, is found 12 nucleotides upstream from the poly(A) tract. Two alternative explanations for these unusual features at the 3' end can be proposed: (1) a deletion of the sequence encoding the C-terminal peptide occurred during the cloning procedure while retaining the sequence corresponding to the poly(A) tract, or (2) the AATAAA sequence at positions 293–298 was recognized as a polyadenylation signal *in vivo* and position 310, which is 12 nucleotides downstream from the AATAAA sequence, served as a polyadenylation site.

The first explanation cannot be excluded although such a cloning artefact has not been described before. The second explanation is a very attractive model for a mechanism regulating rapid changes of Thy-1 gene expression during cell differentiation and is similar to that observed in other systems. For example, during differentiation, B lymphocytes undergo a shift from expression of membrane-bound IgM to secreted IgM. It has been shown that the transcription unit for secreted and membrane-bound IgM heavy chains contains two separate poly(A) addition sites. The basis for selection between these

Received 21 September; accepted 29 October 1982.

1. Reif, A. E., & Allen, J. M. V. *J. exp. Med.* **120**, 413-433 (1964).
2. Douglas, T. C. *J. exp. Med.* **136**, 1054-1062 (1972).
3. Reif, A. E. & Allen, J. M. V. *Nature* **209**, 523 (1966).
4. Stern, P. L. *Nature* **246**, 76-78 (1973).
5. Scheid, M., Boyse, E. A., Carswell, E. A. & Old, L. J. *J. exp. Med.* **135**, 938-955 (1972).
6. Lennon, V. A., Unger, M. & Dulbecco, R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 6093-6097 (1978).
7. Lesley, J. F. & Lennon, V. A. *Nature* **268**, 163-165 (1977).
8. Kuchel, P. W., Campbell, D. G., Barclay, A. N. & Williams, A. F. *Biochem. J.* **169**, 411-417 (1978).
9. Campbell, D. G., Gagnon, J., Reid, B. M. & Williams, A. F., *Biochem. J.* **195**, 15-30 (1981).
10. Efstradiadis, A., Kafatos, F. C., Maxam, A. M. & Maniatis, T. *Cell* **7**, 279-288 (1976).
11. Villa-Komaroff, L. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 3727-3731 (1978).
12. Suggs, S. V., Wallace, R. B., Hirose, T., Kawashima, E. H., & Itakura, K. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6613-6617 (1981).
13. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499-560 (1980).
14. Inouye, M. & Halegoua, S. *CRC Crit. Rev. Biochem.* **7**, 339-371 (1980).
15. Proudfoot, N. J. & Brownless, G. G. *Nature* **263**, 211-214 (1976).
16. Rogers, J. *et al. Cell* **20**, 303-312 (1980).
17. Early, P. *et al., Cell* **20**, 313-319 (1980).
18. Nevins, J. R. & Wilson, N. C. *Nature* **290**, 113-118 (1981).
19. Komuro, K., & Boyse, E. A. *Lancet* **i**, 740-743 (1973).
20. Peterson, P. A., Cunningham, B. A., Berggard, I. & Edelman, G. M. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1697-1701 (1972).
21. Steinmetz, M. *et al. Cell* **24**, 125-134 (1979).
22. Larhammar, D. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 3687-3691 (1982).
23. Williams, A. F., & Gagnon, J. *Science* **216**, 696-703 (1982).
24. Bernard, O., Hozumi, N., & Tonegawa, S. *Cell* **15**, 1133-1144 (1978).
25. Raff, M. C. *Immunology* **19**, 637-650 (1970).
26. Maniatis, T., Fritsch, E. F. & Sambrook, R. *Molecular Cloning*, 115 (Cold Spring Harbor Laboratory, New York, 1982).

Human class II major histocompatibility antigen β -chains are derived from at least three loci

Jan Böhme*, David Owerbach†, Maurizio Denaro*, Åke Lernmark†, Per A. Peterson* & Lars Rask*

* Department of Cell Research, The Wallenberg Laboratory, University of Uppsala, Box 562, S-751 22 Uppsala, Sweden
† Hagedorn Research Laboratory, DK-2820 Gentofte, Denmark

Class II antigens of the major histocompatibility complex (MHC) consist of two glycosylated, membrane-integrated polypeptide chains¹. These cell surface-expressed molecules are involved in several immunobiological events involving cell-cell interactions^{2,3}, most of which seem to require that genetically identical class II antigens, or other molecules controlled by the same region of the MHC, are expressed on the interacting cells⁴. The extensive genetic polymorphism of the class II antigens⁵ has rendered analyses in the human system of the number of non-allelic species of class II antigens difficult, although several laboratories have reported the existence of at least two types of human class II antigens⁶⁻⁹. Here we present the results of experiments using restriction enzyme digestions and separation of DNA from individuals homozygous for the MHC followed by hybridization to human class II antigen α -^{10,11} and β -¹²⁻¹⁴ chain cDNA probes. While the α -chain probe gave only a single hybridization band, the various β -chain probes revealed a more complex pattern that is consistent with the existence of at least three separate β -chain genes or pseudogenes in the human MHC.

The α - and β -chain cDNA probes used in this study are described in Fig. 1 legend. To ascertain that the cDNA probes hybridized only to DNA derived from chromosome 6, on which the human MHC is situated¹⁵, mouse-human cell hybrids¹⁶ were analysed. The *EcoRI*-digested DNA was hybridized in sequence with the four cDNA probes described in Fig. 1 legend. In the stringency conditions used, only DNA from cells containing the human chromosome 6 hybridized with the probes (not shown). To reduce the complexity caused by the extensive polymorphism of the class II antigens this study was carried out using DNA from homozygous typing cells (given by Drs E. Möller and A. Svegaard).

The α -chain cDNA probe corresponds to the DR locus in man¹¹ and is the homologue of the murine I-E/C α -chain¹. The α -chain cDNA probe invariably hybridized to a single band in the stringency conditions used (see Fig. 2). *EcoRI* digests of 15 individual DNA samples, encompassing the haplotypes DR1, DR2, DR3 and DR4, contained in each case a 3.2-kilobase (kb) band which hybridized to the probe (not shown). This is in agreement with previous findings¹⁷ and reinforces the notion that the HLA-DR antigen α -chain shows little or no genetic polymorphism¹⁸. However, in analyses of *BglII*-digested DNA homozygous for DR1, DR3 and DR4, the α -chain cDNA probe hybridized to bands of 3.9, 4.3 and 4.5 kb, respectively (Fig. 2). This pattern requires at least two polymorphic *BglII* sites in or near the DR α -chain gene. One of the previously published DR α -chain cDNA clones has a *BglII* site just 3' of the polyadenylation signal¹⁹ which is absent from other clones^{11,17}. Two possibilities arise if this polymorphism contributes to the pattern in Fig. 2. First, there may be another polymorphic *BglII* site 3.9 kb in the 5' direction from this site. As the whole gene comprises ~5.5 kb with a 2.3-kb intron between the exons encoding the signal peptide and the first domain, respectively (L. Schenning, unpublished observation) this would indicate a location in the large intron between the two first exons. Second, the other polymorphic *BglII* site may be situated 0.4 kb in the 3' direction to the known one. The observation that untranslated or flanking sequences of the DR α -chain gene contain restriction sites that may be haplotype-specific suggests that genomic blot analyses may be used as a complement in clinical HLA-DR typing.

Figure 2 shows that the α -chain cDNA probe only hybridized to a single band also when increasing amounts of restriction enzyme-cleaved DNA were analysed. This does not mean that the human MHC contains only one α gene as two additional and different α -chain cDNA clones have recently been isolated (ref. 20 and our unpublished results). These α -chain cDNA probes cross-hybridize poorly with the DR α -chain cDNA probe used here. Thus, the hybridization results presented here suggest that the HLA-DR α -chain gene may be analysed separately from other α -chain genes or pseudogenes.

The β -chain cDNA probes used in the hybridization analyses were derived from a cDNA clone distinct from the DR β -chain locus^{1,12-14}. This cDNA clone probably corresponds to the human equivalent of the murine I-A β -chain locus¹². However, in contrast to the DR α -chain cDNA probe, the β -chain cDNA probes (see Fig. 1) seem to cross-hybridize quite extensively with related genes as supported by the homology in their nucleotide sequences¹. In contrast to the HLA-DR α -chains, the β -chains show extensive genetic polymorphism. This is corroborated by the fact that the β_0 cDNA probe (see Fig. 1) gives unique hybridization patterns with DNA homozygous for different MHC haplotypes (Fig. 3A). Regardless of whether the DNA had been digested with *EcoRI* or *BamHI*, six to eight hybridization bands were visualized. Because bands greater than ~6 kb were not resolved as well as the smaller bands, the number of bands is probably an underestimate. However, if the hybridization patterns obtained with the cDNA probe used here adequately reflect the number of β -chain loci, it seems reasonable to conclude that class I antigen heavy-chain genes²¹⁻²³ are at least twice as abundant as the β -chain genes (see below).

Figure 3A shows that the human MHC contains more than one β -chain gene or pseudogene. As the homozygous typing cells, which served as the source of the DNA, were only typed for the HLA-DR locus, it is interesting to note that the β_0 probe gave identical hybridization patterns when the DNAs of two homozygous DR3 samples were analysed (Fig. 3A, lanes *f* and *h*). Similarly, the hybridization patterns of the three independent DR4 DNA samples (Fig. 3A, lanes *i-k*) were identical. This suggests that the various class II antigen β -chain loci are probably in linkage disequilibrium.

Human β -chains seem to be folded into two extracellular immunoglobulin-like domains^{1,24}. Preliminary results suggest

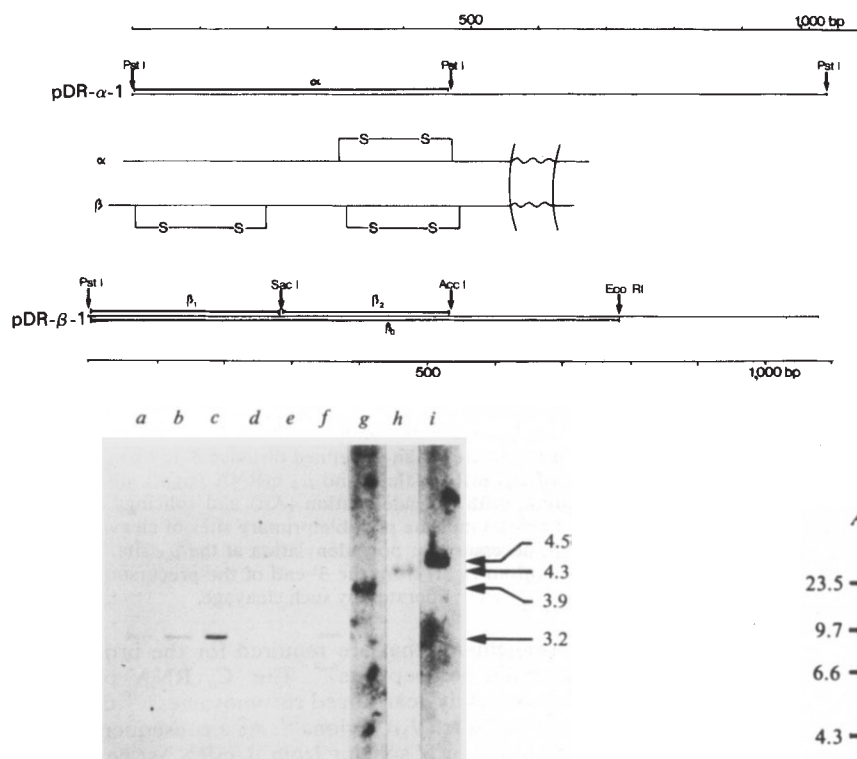


Fig. 2 Characterization of *EcoRI* (a-f) and *BglII* (g-i) genomic restriction fragments by hybridization to the α -chain cDNA probe. Numbers denote molecular weight of hybridizing fragments in kilobases. DNA from homozygous typing cells (BO, DR4, lanes a-c; IES, DR1, lanes d-g; LN, DR3, lane h; BO, lane i) was isolated by solubilization of the cells in 10 mM Tris (pH 7.5), 2 mM EDTA, 10 mM NaCl, 1% SDS and 0.2 mg ml⁻¹ proteinase K and incubated at 37 °C for 16 h. Protein contaminants were removed from high-molecular weight DNA by phenol-chloroform-isoamyl alcohol extraction, and the DNA was recovered by ethanol precipitation. Samples of 5 (lanes a and d), 10 (lanes b, e and g-i) or 20 (lanes c and f) μ g of DNA from each cell line were digested to completion with *EcoRI* and *BglII*, respectively, and electrophoresed overnight on 1% agarose gels in TEAc buffer (40 mM Tris-acetate pH 8.0, 2 mM EDTA) at 20 mA. The DNA in the gels was transferred to nitrocellulose sheets by the method of Southern²⁵. After transfer, the filters were prehybridized (50% formamide, 5 $\times$ SSC, 50 mM PO₄ pH 6.5, 500 μ g ml⁻¹ salmon sperm DNA and 5 $\times$ Denhardt's solution) at 42 °C for 16 h. Radioactively labelled cDNA fragments were added to a concentration of 0.5 $\times$ 10⁶ c.p.m. ml⁻¹ (specific activity >10⁸ c.p.m. μ g⁻¹) in hybridization mixture (50% formamide, 5 $\times$ SSC, 20 mM PO₄ pH 6.5, 200 μ g ml⁻¹ salmon sperm DNA, 1 $\times$ Denhardt's solution and 10% dextran sulphate) and incubated with the filters at 42 °C for 16 h. Then the filters were washed twice in 2 $\times$ SSC, 0.1% SDS at room temperature for 5 min and twice in 0.1 $\times$ SSC, 0.1% SDS at 50 °C for 30 min. The filters were exposed to Kodak XAR-5 film with an intensifying screen at -70 °C.

that each domain, roughly corresponding to the β_1 and β_2 probes, respectively, is encoded in a separate exon (D. Larhammar, unpublished results). Thus, to further analyse the human β -chain loci, hybridization patterns were established using the β_1 and β_2 probes also (see Fig. 1). It was expected that the combined hybridization patterns of the two probes should be very similar to that of the β_0 probe. Figure 3B, which shows the banding patterns of the three probes hybridized to DNA homozygous for DR3 and DR4, respectively, supports this notion. With one exception, the bands hybridizing to the β_0 probe also hybridized to the β_1 and/or the β_2 probes. The 6-kb band, which only hybridized to the β_0 probe, probably contains one or more exons corresponding to the membrane and/or cytoplasmic segments of a β -chain but no other exons. By using the simplified hybridization patterns obtained with the β_1 and

Fig. 1 Construction of the cDNA probes used in the hybridization experiments. In the centre is shown a schematic model of a human class II antigen. The α - and β -chains contain, respectively, one and two disulphide bonds^{11,14}. The wavy lines indicate the membrane-embedded positions of the polypeptide chains which connect the extracellular portions with the cytoplasmic tails. Co-aligned with the polypeptide chains are the pDR- α -1 (top) and pDR- β -1 (bottom) cDNA clones^{11,12}. The cDNA probes used in the hybridization experiments (α , β_0 , β_1 and β_2) are denoted by the solid lines and the restriction enzymes used to excise them from the inserts are shown. After restriction enzyme cleavage, all samples were subjected to preparative polyacrylamide gel electrophoresis followed by electroelution into dialysis tubing.

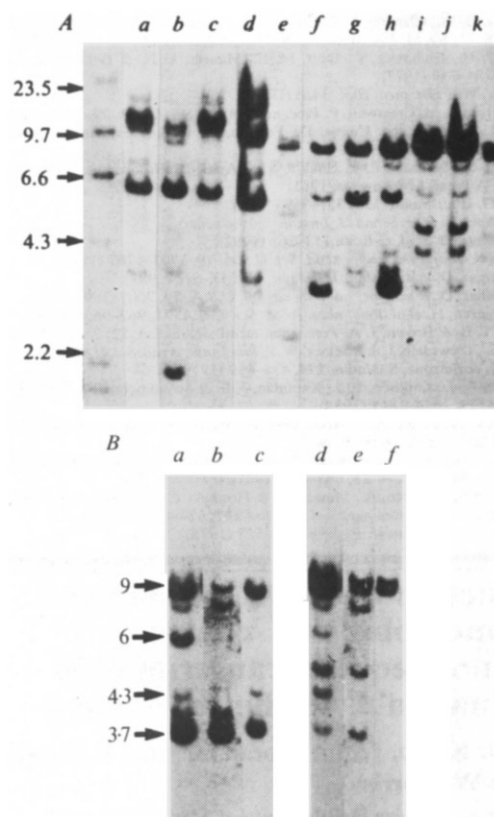


Fig. 3 A, characterization of *EcoRI* (lanes a-d) and *BamHI* (lanes e-k) genomic restriction fragments by hybridization to the β_0 probe. B, characterization of *BamHI* genomic restriction fragments by hybridization to β -domain probes. Numbers denote molecular weight of hybridizing fragments in kilobases.

Methods: A, DNA (10 μ g per lane) from homozygous typing cells (lanes a and e, IES, DR1; lanes b and c, MST, DR2; lanes d and f, LN, DR3; d and k, BO, DR4; h, LK, DR3; i, Leif, DR4; lane j, BB, DR4) was digested, separated, blotted, hybridized, washed and exposed as described in Fig. 2 legend. The values on the left indicate molecular weight of *HindIII*-digested λ marker DNA in kilobases. B, DNA (10 μ g) from homozygous typing cells (lanes a-c, LN, DR3; lanes d-f, BO, DR4) were treated as described in Fig. 2 legend and hybridized to β_0 probe (lanes a and d), β_1 probe (b and e) and β_2 probe (c and f).

β_2 cDNA probes (see Fig. 3B) it should be possible to estimate the minimum number of β -chain genes and pseudogenes present in the human MHC. Thus, the five bands hybridizing to the β_1 probe (Fig. 3B, lane e) can only be interpreted to mean that there are at least three β -chain loci. Of course, this conclusion is based on the reasonable assumption that the β_1 cDNA probe corresponds to a single exon in all types of β -chain genes (see

above). A corollary to this assumption is that a single gene can at most give rise to two fragments hybridizing with the β_1 cDNA probe as the smallest bands comprise several kilobases and the putative exon is smaller than 0.3 kb. The fact that the β_2 cDNA probe hybridizes to three bands (Fig. 3B, lanes c and f) is consistent with the existence of at least three β -chain loci.

Determination of the true numbers of α - and β -chain genes and pseudogenes must await the isolation of these genes and their corresponding mRNAs; such work is in progress in several laboratories. However, our study demonstrates that the class II antigens are derived from a multigene family which is considerably smaller than that noted for the class I antigens.

We thank Mr. Magnus Thörn, who assisted in the early part of this study. M.D. is a holder of a long-term EMBO fellowship. This study was supported by the Swedish Cancer Society and the Gustaf V:s 80-yr fund.

Received 20 September; accepted 5 November 1982.

- Kämpe, O. *et al. Nobel Symp.* Vol. 55 (eds Möller, G. & Möller, E.) (Plenum, New York, in the press).
- Thomas, D. W., Clement, L. & Shevach, E. M. *Immun. Rev.* **40**, 181–204 (1978).
- Miller, S. D., Sy, M. & Claman, H. N. *J. exp. Med.* **145**, 1071–1076 (1977).
- Katz, D. H., Hamaoka, T., Dorf, M. E., Maurer, P. H. & Benacerraf, B. *J. exp. Med.* **138**, 734–739 (1973).
- Bodmer, W. F. *Br. med. Bull.* **34**, 3 (1978).
- Markert, M. L. & Cresswell, P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6101–6104 (1980).
- Tosi, R., Tanigaki, N., Centis, D., Ferrara, G. B. & Pressman, D. *J. exp. Med.* **148**, 1592–1611 (1978).
- Park, M. S., Terasaki, P. I., Nakata, S. & Aoki, D. *Histocompatibility Testing*, 854–860 (UCLA Press, Los Angeles, 1981).
- Corte, G. *et al. Nature* **292**, 357–359 (1981).
- Gustafsson, K. *et al. Scand. J. Immun.* (in the press).
- Larhammar, D. *et al. Cell* **30**, 153–161 (1982).
- Wiman, K. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 1703–1707 (1982).
- Larhammar, D. *et al. Scand. J. Immun.* **14**, 617–622 (1981).
- Larhammar, D. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 3687–3691 (1982).
- van Someren, H. *et al. Proc. natn. Acad. Sci. U.S.A.* **71**, 962–965 (1974).
- Shows, T. B. & Brown, J. A. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2125–2130 (1975).
- Lee, J. S., Trowsdale, J. & Bodmer, W. F. *Proc. natn. Acad. Sci. U.S.A.* **79**, 545–549 (1982).
- Silver, J. & Ferrone, S. *Nature* **279**, 436–439 (1979).
- Korman, A. J., Knudsen, P. J., Kaufman, J. F. & Strominger, J. L. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1844–1848 (1982).
- Auffray, C., Korman, A. J., Roux-Dosseto, M., Bono, R. & Strominger, J. L. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Cami, B., Brégère, F., Abastado, J. P. & Kourilsky, P. *Nature* **291**, 673–675 (1981).
- Steinmetz, M. *et al. Cell* **25**, 683–692 (1981).
- Steinmetz, M., Winoto, A., Minard, K. & Hood, L. *Cell* **28**, 489–498 (1982).
- Kaufman, J. F. & Strominger, J. L. *Nature* **297**, 624–627 (1982).
- Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).

Production of RNA for secreted immunoglobulin μ chains does not require transcriptional termination 5' to the μ_M exons

David J. Kemp, Grant Morahan, Alan F. Cowman & Alan W. Harris

The Walter and Eliza Hall Institute of Medical Research, PO Royal Melbourne Hospital, Victoria 3050, Australia

The mouse immunoglobulin μ gene encodes both secreted and surface-bound μ heavy chains produced by cells of the B lymphoid series. Transcripts of the μ gene are processed into μ mRNA species which differ at their 3' termini, bearing either ' μ_S ' or ' μ_M ' segments, distinguishing secreted and cell-membrane-bound μ polypeptides^{1–3}. During maturation of surface IgM-bearing B cells to IgM-secreting plasma cells, both the total amount of μ mRNA and the ratio of μ_S - to μ_M -terminated mRNA increase greatly^{1,3,4}. Two possible mechanisms for the developmental regulation of 3' RNA processing cannot yet be distinguished. One mechanism would yield the μ_S terminus by specific cleavage of a common precursor transcript encompassing both μ_S and the μ_M exons (Fig. 1), the other by regulated termination of transcription upstream from the μ_M exons^{2,5}. While the first mechanism would produce, as a by-product, RNA fragments containing μ_M exons, the second would not. We report here the detection of such μ_M fragments in cells producing predominantly μ_S -terminated RNA species.

Previously, we found that in some T lymphoma and myeloid tumour lines the C_μ genes are transcribed in the absence of the

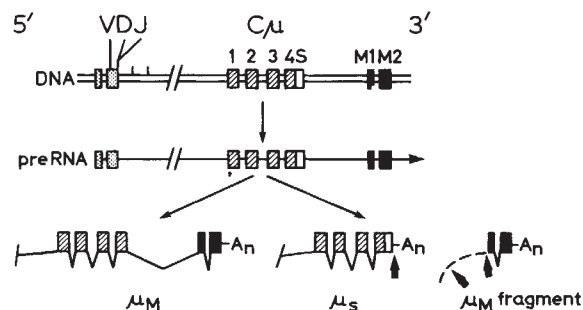


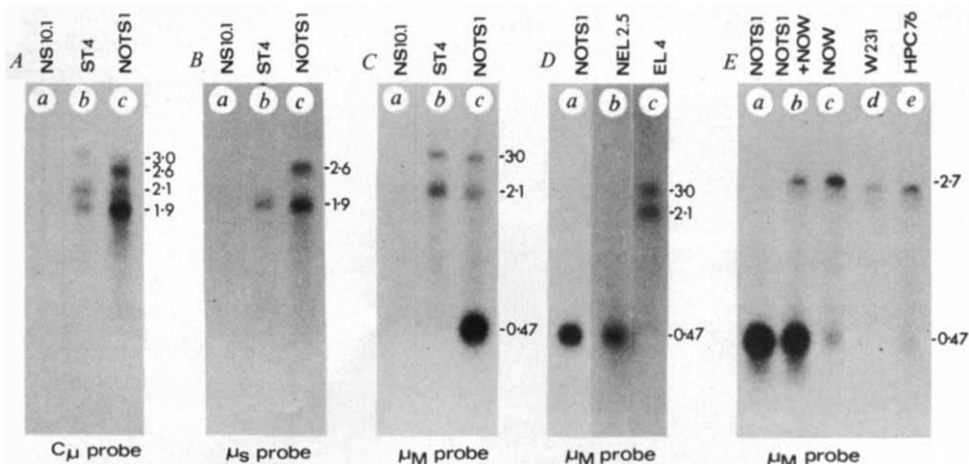
Fig. 1 Model for generation of μ_S and μ_M 3' termini from a common precursor. A productively rearranged C_μ gene bearing V_H , D_H and J_H segments is shown (top). The precursor RNA (middle) includes the exons (boxes) encoding V_H - D_H - J_H , C_μ exons 1–4, μ_S which is contiguous with C_μ , and the two μ_M exons; it terminates at an undefined distance 3' to exon μ_{M2} . The 3' ends of μ_M mRNA (left) and μ_S mRNA (right) are shown at the bottom, with polyadenylation (An) and splittings indicated. Upward arrows indicate possible primary sites of cleavage of the transcript necessary for polyadenylation at the μ_S site. ' μ_M fragments' (bottom right) from the 3' end of the precursor would be liberated by such cleavage.

gene rearrangements that are required for the production of immunoglobulin polypeptides^{6–8}. The ' C_μ RNA' products of non- or incompletely-rearranged chromosomes^{6–10} do not bear heavy chain variable (V_H) regions^{7,9}. As a consequence, the C_μ RNAs differ in their 5' splicing from μ mRNAs; nevertheless, they are all terminated by either μ_S or μ_M at the 3' end^{7,9}. The functionally rearranged μ genes of B cells, when present together with a plasmacytoma genome in a cell fusion hybrid, produce highly elevated ratios of μ_S : μ_M mRNA, a large increase in total μ mRNA ('transcriptional amplification') and consequently, greatly increased amounts of μ polypeptides for secretion^{1–5,11}. Hence, developmental control of both transcriptional amplification and 3' processing are most probably mediated by *trans*-acting regulatory molecules present in plasmacytoma cells. To test whether expression of unrearranged (or incompletely rearranged) μ genes is subject to similar control, we constructed hybrid cell lines NOTS1 and NEL2.5 by fusion of C_μ RNA-producing T lymphoma lines ST4 and EL4, respectively, with plasmacytoma line NS1 (Fig. 2A–C) and examined the hybrids for μ gene transcripts. Because NS1¹² is a descendant of the MOPC 21 tumour, known to lack all μ genes¹³, any μ RNA in these hybrids must be transcribed from T lymphoma μ genes. Both μ alleles in ST4 have undergone D- J_H joining, whilst both μ alleles in EL4 remain in the unrearranged germ-line condition (S. Gerondakis, O. Bernard, S. Cory and J. M. Adams, personal communication).

Polyadenylated RNA species bearing C_μ , μ_S and μ_M sequences in somatic hybrids and parental cells were characterized and approximately quantified by electrophoresis and transfer to a filter¹⁴, followed by hybridization with probes derived from cloned segments of the μ gene region as shown in Fig. 2A. Unlike the 100-fold amplification of μ mRNA in B-cell $\times$ plasmacytoma hybrids, the present hybrids showed only a two- to fourfold increase in the total amounts of C_μ gene transcripts (Fig. 2A, tracks b, c). Nevertheless, the hybrids continued to produce μ mRNA at the same high level as NS1, about 50–100-fold greater than the amount of μ RNA (data not shown). Thus, unlike transcriptional competence which does not require J_H rearrangement, transcriptional amplification apparently requires translocated V_H sequences. A simple interpretation is that the regulatory molecules from NS1 act on sequences located in or 5' to translocated V_H genes.

The major differences between the T lymphoma parent ST4 and its derived hybrid NOTS1 were a 4- and 10-fold increase of the 1.9- and 2.6-kilobase (kb) μ_S -bearing C_μ RNA species of NOTS1 relative to the 2.1- and 3.0-kb μ_M -bearing species (Fig. 2B, tracks b, c). Further, the hybrid cells contained an abundant small RNA bearing μ_M (Fig. 2C, track c) but not C_μ

Fig. 2 A–C, C_{μ} , μ_S - and μ_M -terminated C_{μ} RNAs in T lymphoma-plasmacytoma hybrid line NOTS1 and its parental lines. Cells were grown as described previously⁶. NS1 0.1, a ouabain-resistant clone of P3NS1/1-Ag-4-1 (NS1), was selected by growth in medium containing $10^{-3.3}$ M ouabain²¹ and 10^7 cells were fused with an equal number of STRij-4-2.2 (ST4)⁶ cells by the polyethylene glycol method²². The cloned hybrid cell line NOTS1 was selected by growth in selective medium²³ containing hypoxanthine, aminopterin, thymidine (HAT) and ouabain. Total cellular poly(A)⁺ RNA was prepared¹⁴, and samples (5 μ g) fractionated



by electrophoresis on a 1.5% agarose gel in the presence of 5 mM methylmercuric hydroxide²⁴, transferred to diazobenzylxymethyl paper¹⁴ and hybridized with nick-translated probes (specific activity $1-2 \times 10^8$ c.p.m. per μ g) as described previously⁶. The C_{μ} probe (A) was a *PvuII*-*Bam*HI fragment of cDNA clone pH76 μ 17²⁵, comprising part of exon $C_{\mu 2}$. The μ_S probe (B) was a *Hha*-*Hind*III fragment of *Hind*III subclone $C_{\mu Hb}$ (ref. 7), derived from chromosomal clone Ch.H76 μ 126. The μ_M probe (C) was a *Kpn*-*Hind*III fragment of *Hind*III subclone $C_{\mu Hc}$ (ref. 7) comprising both μ_M exons. These fragments were purified from the relevant subclones by electrophoresis on 5% polyacrylamide gels. Sizes are shown in kilobases. D, μ_M RNA fragments in two independent T lymphoma-plasmacytoma hybrid cell lines. Hybrid cell line NEL2.5 was selected after fusion of NS1 and EL-4BU.2²⁷ by growth in HAT medium. Details as in A, except that a 2.5% gel was used, and *Hae*III fragments of Φ X174 DNA were used as markers. E, μ_M RNA fragments in cells of the B lineage. Hybrid cell line NOW10-12 was selected after fusion of NS1 0.1 and a bromodeoxyuridine-resistant mutant of the B lymphoma line WEHI-231²⁸ by growth in HAT plus ouabain as above. Other details were as in A. Track b contained a mixture of equal amounts of RNA from NOTS1 and NOW10-12.

(Fig. 2A) or μ_S (Fig. 2B) sequences. This latter ' μ_M fragment' could not be found in the parental lines, nor in several other T- and B-cell lines tested (18-81, RAW8, WEHI-222, data not shown), but was readily detectable in the NEL2.5 hybrid as well as in NOTS1 (Fig. 2D). Its size was 470 ± 50 bases (Fig. 2D), sufficient to contain just exons μ_{M1} and μ_{M2} with a 75-base poly(A)tail^{1,2}. By densitometry of autoradiographs such as that shown in track c of Fig. 2A and C, we estimated that there are about 2 moles of the μ_M fragment per mole of the total (that is 1.9 plus 2.6 kb) μ_S -terminated species in NOTS1, using the 2.1- and 3.9-kb species as internal standards. Allowing for the higher efficiency of detection of small RNA species¹⁴, the total amount of μ_M fragment is therefore similar to that of μ_S termini. The appearance of approximately equimolar amounts of the μ_M fragment accompanying the increase in amounts of μ_S -terminated C_{μ} RNA species in the hybrid cells relative to the T lymphoma parents suggests that the fragment is a by-product in the formation of μ_S -terminated C_{μ} RNA from precursor transcripts containing both μ_S and μ_M . It was therefore important to determine whether the fragment could be detected in cells producing high ratios of functional μ_S mRNA.

To test this we constructed hybrid NOW10-12 by fusion of B lymphoma line WEHI-231 (W231) with plasmacytoma line NS1 (Fig. 2E). NOW10-12 secreted high levels of IgM and contained 50-100-fold more μ_S mRNA than W231 (data not shown), in agreement with other work^{1,3,4}. A μ_M fragment of much lower abundance but of identical size to that in NOTS1 was also evident in NOW10-12 (Fig. 2E, tracks a-c), but not in W231 (track d). Similarly, the μ_M fragment was detectable in IgM-secreting plasmacytoma HPC 76 (Fig. 2E, track e).

To examine the structure of the μ_M fragment, we used a cDNA clone (pST4 μ 105) constructed from ST4 C_{μ} RNA. This clone contains exons $C_{\mu 4}$, μ_{M1} and μ_{M2} (see Fig. 4a). A probe located to the right of the *Xba* site in pST4 μ 105 hybridized to the μ_M fragment (data not shown), demonstrating that this fragment contains sequences from the 3' end of μ_{M2} . We used probes from both the cDNA and chromosomal⁸ μ_M clones to determine the structure of the μ_M fragments of NOTS1 more precisely, measuring the lengths of hybrids by the nuclease method of Berk and Sharp¹⁵. The results and our interpretation of them are presented in Figs 3 and 4.

The μ_M fragment exists in two forms (Fig. 4c). In both forms exon μ_{M1} is correctly spliced to μ_{M2} , but μ_{M1} is not spliced to exon $C_{\mu 4}$. The minor but longer form (μ_{MB}) results from

cleavage at the 5' boundary of exon μ_{M1} , while the major, shorter form (μ_{MA}) is cleaved at a site 20-24 bases further downstream. Both forms may simply be cleaved at these sites or could be spliced to very short segments of unknown origin which we would be unable to detect.

The sequence at the 5' boundary of the μ_{M1} exon^{1,2}, 5'-AG/AGGGGGAGGTGAATGCTGAG/GAG/GAAGCC-3' (where the slash after the first AG denotes the μ_{M1} splice boundary), has several features which may be relevant to cleavage at the μ_{MA} boundary. First, the sequence beginning 18 bases 3' to the splice boundary is identical to that beginning at base 1 in 8 out of 10 positions (indicated by underlining), and purines occupy all these positions. Second, two additional potential splice sites (indicated by slashes) are located 20 and 23 bases 3' to the splice boundary, agreeing with the consensus splice sequence (Y-YYY-CAG/)¹⁶ at four and three of seven positions, respectively. Both are out of phase with respect to both the repeated sequence at position 1 and the reading frame of exon $C_{\mu 4}$, and would lead to termination after four and three amino acids, respectively, if spliced to $C_{\mu 4}$ in an otherwise functional μ mRNA. The 5' boundary of the μ_{MB} fragment is probably generated by 'slicing', that is, cleavage at the μ_{M1} splice site without subsequent ligation. Similarly, the μ_{MA} fragment may be generated by slicing at a 'cryptic' splice site (of which other examples are known¹⁷⁻¹⁹), revealed only because normal splicing cannot occur.

The detection of μ_M RNA fragments in cells which produce predominantly μ_S -terminated μ RNAs suggests that the fragments are derived as a by-product of μ_S production from a precursor which extends through the μ_M exons. We conclude that specific termination of transcription 5' to the μ_M exons is not required for μ_S production—hence a specific cleavage model is implicated^{2,5}. Several forms of the cleavage model remain to be distinguished. For example, it remains unclear whether cleavage of a common full-length μ_S , μ_M precursor occurs as shown in Fig. 1, or whether cleavage can occur before the polymerase reaches the μ_M exons. The primary cleavage may well occur at the μ_S poly(A) addition site. However, cleavage at any other site 3' to this could prevent μ_M production and allow poly(A) addition at the μ_S site. It is evident from the present data that there are large differences in the abundance of μ_M fragments in the T hybrids and plasmacytoma HPC76 and this could arise because the C_{μ} RNA transcripts differ structurally from μ mRNA transcripts, because transcription

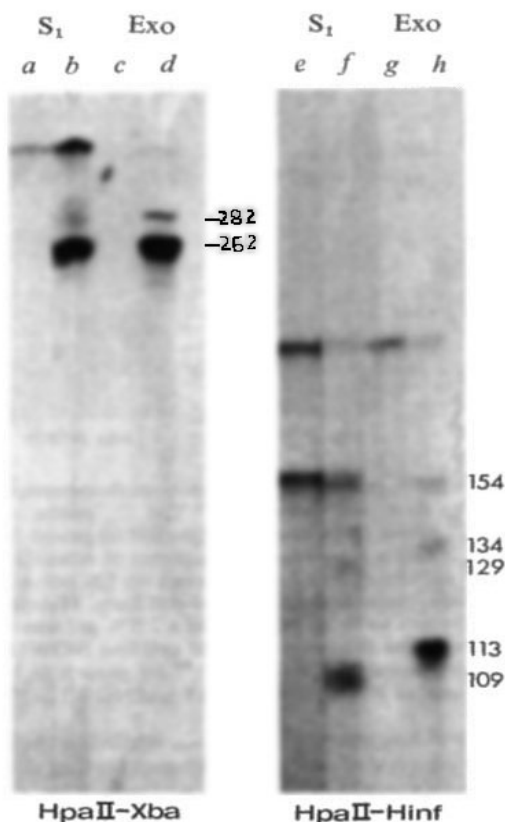


Fig. 3 Analysis of the μ_M RNA fragment by nuclease mapping¹⁵. Labelled DNA fragments were incubated in a buffered solution containing 60% formamide and 0.42 M NaCl for 24 h at 60 °C with (tracks b, d, f, h) or without (tracks a, c, e, g) $\sim 5 \mu\text{g}$ NOTS1 RNA (purified by sucrose gradient centrifugation to remove the larger C_μ RNAs from the μ_M fragment, as established by Northern analysis¹⁴). After digestion with nuclease S₁ (tracks a, b, e, f) or exonuclease VII (tracks c, d, g, h), aliquots were fractionated on 0.2 mm 6% acrylamide-urea gels and undigested DNA fragments detected by autoradiography. Sizes, indicated in bases, were determined by reference to end-labelled pBR322 *Hinf* + *Rsa* markers. For tracks a–d, the probe was an *HpaII*–*XbaI* fragment of pST4 μ 105 spanning the C_μ 4–M1 and M1–M2 boundaries (Fig. 4a), labelled by replacement synthesis with T4 DNA polymerase²⁹. For tracks e–h, the labelled *HpaII*–*Xba* fragment was cut with *HinfI*.

continues through the μ_M exons more commonly in the T hybrids or because the μ_M fragments are more stable in the T hybrids. Regardless of these alternatives, our data are incompatible with the hypothesis that specific termination of transcription 5' to the μ_M exons is necessary for μ_S production.

The C_μ gene is a segment of a transcription unit arranged ($V_H D_H J_H$)– C_μ – C_δ , which can simultaneously produce μ_S , μ_M , δ_S and δ_M mRNAs. In early B cells producing only μ chains,

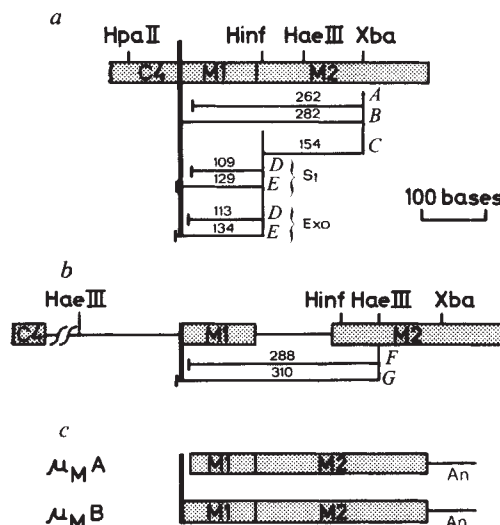


Fig. 4 Structure of the μ_M RNA fragments. a, cDNA clone pST4 μ 105 was constructed by synthesizing double-stranded DNA from polyadenylated RNA of ST4 and inserting this into the *Pst* site of pBR322 by the dG:dC tailing procedure³⁰. $\sim 10^5$ clones were screened with a chromosomal probe derived from $C_{\mu Hc}$ (ref. 7). The positions of *BglII*, *HincII*, *Hinf*, *HpaII*, *PvuII* and *XbaI* restriction sites were shown to be identical to those in cDNA clone M6¹; only the relevant sites are indicated. Exons C_μ 4, μ_{M1} and μ_{M2} are shown. The left edge of the thick vertical line defines the C_μ 4– μ_{M1} boundary. Below the map, fragments a–e observed by S₁ and exonuclease VII analysis in Fig. 3, are aligned with the map. b, Chromosomal DNA clone $C_{\mu Hc}$ (ref. 7). The introns separating C_μ 4, μ_{M1} and μ_{M1} – μ_{M2} are shown as thin lines, the exons as stippled bars. The *HaeIII* fragment shown was labelled by replacement synthesis with T4 DNA polymerase²⁹, hybridized to the purified μ_M fragment and digested with exonuclease VII. Fragments F and G of 288 bases and 310 bases, respectively, were produced (data not shown). Hence, fragments A–E are related to the cDNA map as shown above, since the μ_M RNA fragments are contiguous with the cDNA in the region of the 3' *HaeIII* site. The locations of the 5' ends of DNA fragments F and G must therefore be at this *HaeIII* site. Hence, the 5' ends of the μ_M RNA fragments do not extend into the intervening sequence 5' to μ_{M1} . c, Structure of the μ_M fragments, as deduced from the sizes of nuclease fragments A–G above. An designates poly(A).

the C_δ gene is methylated and transcriptionally inactive, implying that C_δ production is controlled by specific termination of transcription²⁰. The results here suggest that, in contrast, differential production of surface and secreted immunoglobulin may be determined by differential cleavage of common precursor transcripts.

We thank Kirstin Easton for technical assistance, and Brett Tyler, Jerry Adams and Suzanne Cory for helpful discussions and gifts of DNA. This work was supported by grants from the NC1 (USA) (1R01-CA29634-01), the NH and MRC, (Australia), the Drakensberg Trust and, for G.M. and A.F.C., Commonwealth Postgraduate Research Awards.

Received 25 October; accepted 1 November 1982.

- Rogers, J. *et al. Cell* **20**, 303–312 (1980).
- Early, P. *et al. Cell* **20**, 313–319 (1980).
- Alt, F. W. *et al. Cell* **20**, 293–301 (1980).
- Perry, R. P., Kelley, D. E., Coleclough, C. & Kearney, J. F. *Proc. natn. Acad. Sci. U.S.A.* **78**, 247–251 (1981).
- Rogers, J. *et al. Cell* **26**, 19–27 (1981).
- Kemp, D. J., Harris, A. W., Cory, S. & Adams, J. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2876–2880 (1980).
- Kemp, D. J., Harris, A. W. & Adams, J. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7400–7404 (1980).
- Cory, S., Adams, J. M. & Kemp, D. J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4627–4631 (1980).
- Alt, F. W., Rosenberg, N., Enea, V., Siden, E. & Baltimore, D. *Molec. cell. Biol.* **2**, 386–400 (1982).
- Zuniga, M.-C., D. Eustachio, P. & Ruddie, N. H. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3015–3019 (1982).
- Raschke, W. C., Mather, E. L. & Koshland, M. E. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3469–3472 (1979).
- Köhler, G., Howe, S. C. & Milstein, C. *Eur. J. Immun.* **6**, 292–295 (1976).
- Cory, S. & Adams, J. M. *Cell* **19**, 37–51 (1980).

- Alwine, J. C., Kemp, D. J. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5350–5354 (1977).
- Berk, A. J. & Sharp, P. A. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1274–1278 (1978).
- Sharp, P. A. *Cell* **23**, 643–646 (1981).
- Dunnick, W., Rabbitts, T. H. & Milstein, C. *Nature* **286**, 669–675 (1980).
- Seidman, J. G. & Leder, P. *Nature* **286**, 779–783 (1980).
- Bothwell, A. L. M. *et al. Nature* **290**, 65–67 (1981).
- Rogers, J. & Wall, R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7497–7501 (1981).
- Baker, R. M. *et al. Cell* **1**, 9–21 (1974).
- Galfre, G., Howe, S. C., Milstein, C., Butcher, G. W. & Howard, J. C. *Nature* **266**, 550–552 (1977).
- Littlefield, J. W. *Science* **145**, 709–710 (1964).
- Bailey, J. M. & Davidson, N. *Analyt. Biochem.* **70**, 75–85 (1976).
- Adams, J. M., Gough, N. M., Webb, E. A., Jackson, J. & Cory, S. *Biochemistry* **19**, 2711–2719 (1980).
- Gough, N. M., Kemp, D. J., Tyler, B. M., Adams, J. M. & Cory, S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 554–558 (1980).
- Mohit, B. & Fan, K. *Science* **171**, 75–77 (1971).
- Warner, N. L., Daley, M. J., Richey, J. & Spellman, C. *Immun. Rev.* **48**, 197–243 (1979).
- O'Farrell, P. H., Kuttner, E. & Nakanishi, M. *Molec. gen. Genet.* **179**, 421–435 (1980).
- Bolivar, F. *et al. Gene* **2**, 95–113 (1977).

A mouse type 2 *Alu* sequence (M2) is mobile in the genome

Ryo Kominami & Masami Muramatsu

Department of Biochemistry, Cancer Institute, Japanese Foundation for Cancer Research, 1-37-1, Kami-Ikebukuro, Toshima-ku, Tokyo 170, Japan

Kazuo Moriwaki

Department of Cytogenetics, National Institute of Genetics, Tanida, Mishima 411, Japan

The *Alu* and its equivalent families of interspersed repetitive DNA sequences have been found in various mammalian genomes. It has been proposed that some of them might move around the genome like known prokaryotic and eukaryotic transposable elements, as most of these sequences are flanked by short direct repeats at both ends¹⁻⁷. To prove that this is the case, however, one must demonstrate the existence of homologous sequences of DNA with and without *Alu* insertion among the genomes of different strains or individuals of a species. While studying a polymorphic repetitive sequence (PR1) originally found in the spacer region of mouse ribosomal RNA genes⁸, we have now found that a sequence similar to the CHO type 2 *Alu*-equivalent element⁷, designated M2, is inserted within a PR1 sequence which is located outside the ribosomal RNA gene and that this M2 segment is flanked by a short direct repeat at both ends. Furthermore, this PR1 segment containing M2 is detected only in the BALB/c strain among the laboratory mice and wild mouse subspecies examined. These facts suggest that the M2 sequence has been inserted into PR1 sequence relatively recently during evolution of mouse strains and support the idea that at least some of the *Alu*-equivalent families are mobile in the genome. Recently, Grimaldi and Singer⁹ reported an African green monkey α -satellite sequence that was interrupted by an *Alu* element.

Several kinds of repetitive sequence exist in the spacer region of mouse ribosomal RNA genes^{8,10,11}. We have investigated one that was found near the 3' end of 28S ribosomal RNA. The probe of *Eco*RI 6.6-kilobase (kb) rDNA fragment containing the 3' end of 28S rRNA gene (Fig. 1) revealed several bands in addition to the 6.6-kb band when hybridized to a Southern blot of mouse genomic DNA digested with *Eco*RI, indicating that a repetitive sequence(s) is present in the spacer region covered by this DNA fragment. We have isolated a clone, hybridizing to the repetitive sequence, carrying *Eco*RI 3.5-kb DNA, which corresponds to the same-sized band in the *Eco*RI digest of genomic DNA (see Fig. 3a). This repetitive sequence is designated as PR1, the details of which will be published elsewhere (R.K. *et al.*).

The restriction endonuclease maps of the 6.6-kb rDNA and the 3.5-kb DNA are shown in Fig. 1. PR1 regions were determined by cross-hybridization and partial sequence analysis (data not shown). The homologous region was separated into at least three segments. The homologous regions in each clone are illustrated by bold lines. Figure 2 compares the nucleotide sequences of the homologous and adjacent regions of the two cloned fragments. One of the homologous regions in the 3.5-kb DNA, consisting of 70 base pairs (bp), is interrupted by a segment which shows homology with the CHO type 2 *Alu*-equivalent sequence reported by Hynes and Jelinek⁷. This segment, designated as M2 (abbreviated from mouse type 2 *Alu* equivalent), contains oligo(dA) at one end and is flanked by 10-bp direct repeats. The segment between the one direct repeat and the opposite oligo(dA) tract consists of 178 nucleotides. The presence of direct repeats at the ends of the inserted M2 segment suggests that they are formed by a target site duplication.

In prokaryotes and eukaryotes, the insertion of a transposable element leads to a duplication of a small number of nucleotides

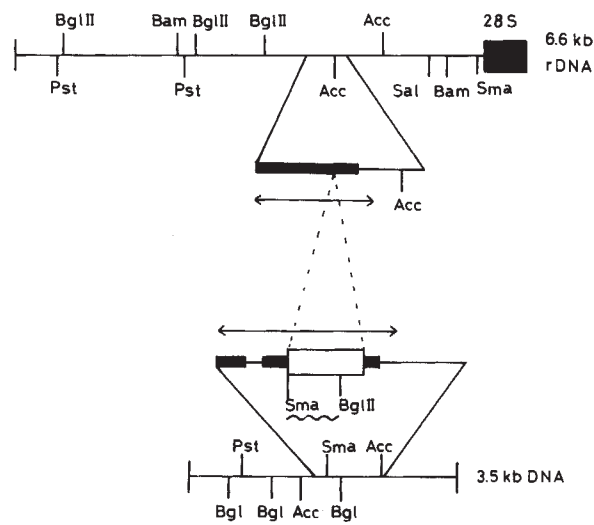


Fig. 1 Maps of 6.6-kb *Eco*RI fragment of rDNA and cloned 3.5-kb *Eco*RI fragment which has homology with the 6.6-kb fragment. A clone containing *Eco*RI 3.5-kb DNA was isolated from a BALB/c mouse genomic library as a clone hybridizing to nick-translated 6.6-kb rDNA described here. Vertical lines indicate specified restriction endonuclease cleavage sites. Solid box shows a coding region of the 3' end of 28S ribosomal RNA¹⁸ and open box a mouse type 2 (M2) interspersed repetitive sequence. Bold lines define the regions that are homologous to each other (part of PR1 sequence). The segments indicated by wavy lines are those used as probes in the Southern blot hybridization experiments described in Fig. 3. The regions whose nucleotide sequences are shown in Fig. 2 are indicated by arrows.

```

3.5kb  GCCATCTTACTGTAACCTTCATGCATTTCAATGAGTAT (GA)nTCGCTGA
6.6kb  .....CT.....C'.....C'G'..... 'A'.....
          TCCCCAGCTCCACAAACACCCAAATC TAAACAAAGAAACCTCCCGG
          .....C'.....GG'.....
          CTGGTGAGATGGTTCAGTGGGTAGAGCACCCGACTGCTCTCCGAAGGTC
          CAGAGTTCAAATCCAGCAACCATGGTGGCTCACCAACCATCCGTAAC
          AAGATCTGACTCCCTCTTCTGGAGTGCTGAAGACAGCTACAGTGTACTT
          ACATATAATAATAATAAATCTTTAAAAAAGAAACCTCCCAACCA
          .....
          GTCAACTAATTTCAATGAATAAATGAACAGGGTCAACTAGTTTGATAT
          'C'.....'GGTCAACTA'TT'C'.....TGAAT'A'CGA'ACAG'CCA
  
```

Fig. 2 Sequences of the boundary regions of a mouse type 2 *Alu* element (M2) and of the homologous segment of the ribosomal DNA. The first line shows the sequence of 3.5-kb DNA. The line below indicates the homologous region in rDNA, the sequence of which is marked by a dot where it is identical with that of 3.5-kb DNA. The short arrows overlying the sequence indicate the direct repeats that flank the extremities of the M2 sequence. (GA)_n indicates the nucleotide sequence composed of 60-fold alternating (GA) clusters with some irregularity. Nucleotide sequences were determined by the method of Maxam and Gilbert¹⁹.

at the target site^{12,13}. Therefore, the repetition present at both sides of the M2 sequence indicates the transposable nature of this element. These *Alu*-equivalent families, however, show important differences in several properties from known transposable elements. For instance, *Alu* sequences consist of much shorter and more divergent sequences than the transposable or insertion elements. These latter elements carry long terminal repeats. Only, a *Drosophila melanogaster* mobile element called

Fig. 3 Blotting patterns of genomic DNAs from laboratory and wild mice hybridized with the sequence flanking the M2 element (a) or M2 element itself (b). Each track contains the DNA from livers of the following mice. 1, *Mus musculus molossinus*, caught at Momoyama, in central Japan. 2, *M. m. molossinus*, caught at Hokkaido. 3, *Mus musculus bactrianus*, caught in Afghanistan. 4, *Mus musculus domesticus*, caught at Windsor, Canada. 5, BALB/c; 6, C3H/He; 7, C57BL; 8, A/J; 9, SM/J. Livers were excised from mice and homogenized in a Potter homogenizer with 20 vol of 0.02 M Tris-HCl, pH 7.6, 0.1 M NaCl, 2 mM MgCl₂ at 0°C. Nuclei were pelleted with a low speed centrifugation and resuspended in 20 vol of 50 mM Tris-HCl, pH 7.6, 5 mM EDTA, 0.1 M NaCl and 0.5% SDS containing 200 µg ml⁻¹ proteinase K (Merk). The viscous lysate was digested for 15 h and extracted with phenol three times. The aqueous layer was dialysed extensively against 0.05 M Tris-HCl, pH 7.6, 5 mM EDTA, 0.01 M NaCl at 4°C, and finally against 0.01 M Tris-HCl, pH 7.6, 0.1 mM EDTA²⁰. The DNA was digested at a concentration of 50 µg ml⁻¹ with *EcoRI* (1 unit enzyme per µg of DNA, Takara Shuzo, Japan) in 50 mM Tris-HCl, pH 7.6, 0.1 M NaCl, 5 mM MgCl₂, and 1 mM dithiothreitol at 37°C overnight. Digests (5 µg of DNA) were electrophoresed on 0.8% agarose gels in 40 mM Tris-acetate, pH 7.8, 5 mM sodium acetate, 5 mM EDTA at 5 V cm⁻¹ for 5 h. After electrophoresis, the DNA was transferred to nitrocellulose filters by the method of Southern²¹. The *PstI-SmaI* 1.2-kb fragment adjacent to the M2 element (a) and the *SmaI-BglII* 112-bp fragment (b) shown in Fig. 1 were nick-translated with [α -³²P]-labelled dCTP to a specific activity of 10⁸ c.p.m. per µg. Filters were incubated in 0.5 ml of hybridization mixture containing 3 × 10⁶ c.p.m. of ³²P-labelled cloned DNA after denaturation at 100°C for 10 min. Hybridization was done in 50% formamide, 0.04M PIPES-NaOH, pH 6.8, 0.5 M NaCl, 0.1% SDS, 1 mM EDTA and 25 µg ml⁻¹ of denatured *Escherichia coli* DNA containing 10×Denhardt's solution²² at 40°C overnight. After hybridization, filters were washed extensively in 0.1×SSC, dried and autoradiographed for 2 h with intensifier screens. The approximate sizes of DNA in kilobase pairs (kb) in the bands are indicated.

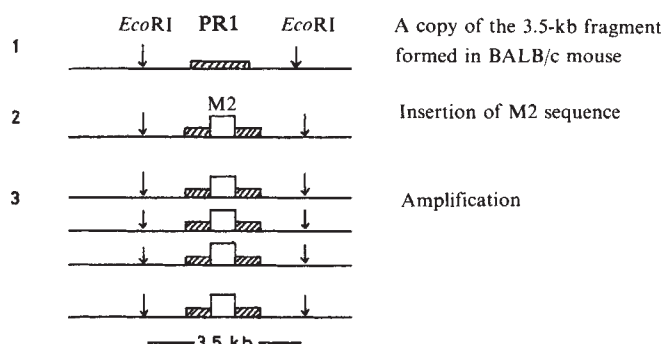
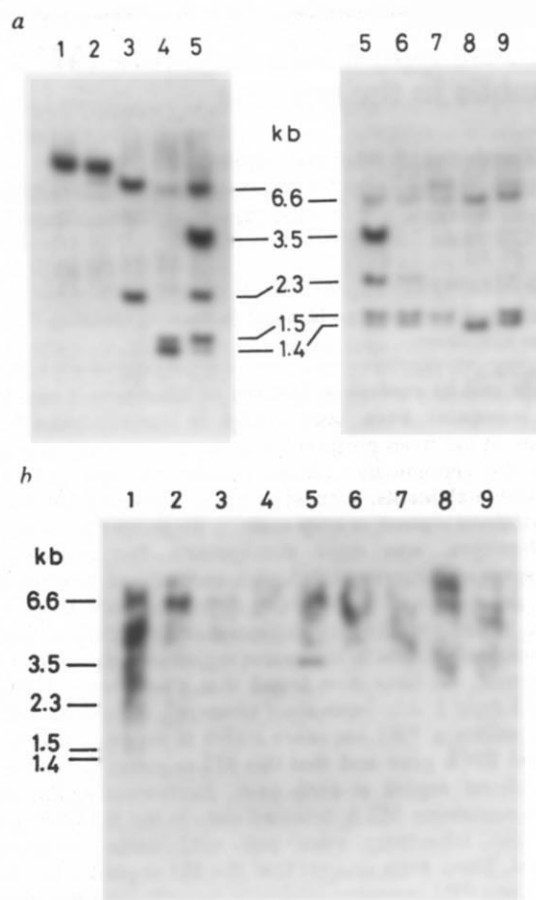


Fig. 4 A model for the generation of PR1 inserted by a type 2 *Alu* sequence (M2) in 3.5-kb *EcoRI* fragment. The details are explained in the text.

101F is similar to *Alu* family members although it is much longer¹⁴. The members of this element found in the chromocentre show extensive site variability between stocks of *D. melanogaster*, demonstrating that this element is mobile in the genome.

The movement of the M2 element, if it occurs, should result in a change of genomic organization. To investigate the arrangement of M2 and the flanking sequences during evolution, we examined several strains of laboratory mice and wild mouse subspecies. DNA from each mouse was digested with *EcoRI* and subjected to Southern hybridization using the sequence (*SmaI-PstI* 1.2 kb DNA) flanking the M2 element as probe (Fig. 3a). The probe consists of most of the PR1 sequence. The DNA samples revealed patterns with substantial polymorphism. The 6.6-kb rDNA band and the other bands obtained here comprised several hundred repetitive segments per haploid genome as shown by a DNA dosage experiment (data not shown), forming polymorphic repetitive sequence families containing PR1 sequence (termed PR1 family).

Note that the 3.5-kb band containing the M2 element was detected only in BALB/c mouse, not in other mice. Although

this 3.5-kb band comprised ~300 copies in the genome, even a single copy DNA band could have been detected in the conditions used. The data indicate that not a single copy of this member exists in other mice. DNAs from different individuals of BALB/c strain and from other tissues (kidney and spleen) did not show any difference in bands as far as was examined (data not shown).

The M2 probe (*SmaI-BglII* 112 bp DNA) showed only one conspicuous band of 3.5 kb in BALB/c DNA but mostly smears in other laboratory strains (Fig. 3), suggesting that the other bands hybridizing to PR1 sequence (Fig. 3a) did not contain the M2 element sequenced here. The PR1 families visualized as satellite-like bands in Southern blots seem to have been created by recombination of the original PR1 sequence with other DNA segments followed by subsequent amplification. Apparently, the segment covering *EcoRI* 3.5-kb DNA detected in BALB/c mice was generated very recently in mouse evolution, at least after the BALB/c strain or its ancestor had separated from other strains of laboratory mouse. The M2 element was inserted in one of the PR1 segments but not in the others during the course (Fig. 4). Thus, this element must have moved into a place newly formed in the BALB/c mouse genome but not in the PR1 sequences of the other mouse strains and subspecies examined here.

These findings together strongly support the conclusion that some members of the M2 family are mobile. The mechanism of transposition of *Alu*-equivalent families remains to be clarified, although several models have recently been proposed^{7,15,16}.

Another putative transposable element having a structure similar to a retrovirus long terminal repeat has been reported in the bovine genome¹⁷.

We thank Drs T. Kataoka and T. Honjo for their gift of a BALB/c *EcoRI* partial library and Drs Y. Urano, N. Okada, Y. Oshima, and H. Yoshikura for helpful comments on the manuscript. These studies were supported in part by grants from the Ministry of Education, Science and Culture of Japan.

Received 31 August; accepted 26 October 1982.

1. Houck, C. M., Rinchart, F. F. & Schmid, C. W. *J. molec. Biol.* **132**, 289–306 (1979).
2. Jelinek, W. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 1389–1402 (1980).
3. Krayev, A. S. *et al. Nucleic Acids Res.* **8**, 1201–1215 (1980).
4. Dhruva, B. R., Shenk, T. & Subramaniam, K. N. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4514–4518 (1980).
5. Elder, J. T., Pan, J., Duncan, C. H. & Weissman, S. M. *Nucleic Acids Res.* **9**, 1171–1189 (1981).
6. Schmid, C. W. & Jelinek, W. R. *Science* **216**, 1065–1070 (1982).
7. Haynes, S. R., Toomey, T. P., Leinwand, L. & Jelinek, W. R. *Molec. cell. Biol.* **1**, 573–583 (1981).
8. Kominami, R., Urano, Y., Mishima, Y. & Muramatsu, M. *Nucleic Acids Res.* **9**, 3219–3233 (1981).
9. Grimaldi, G. & Singer, M. F. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1497–1500 (1982).

10. Arnheim, N. *et al. Cell* **22**, 179–185 (1980).
11. Miesfeld, R., Krystal, M. & Arnheim, N. *Nucleic Acids Res.* **9**, 5981–5947 (1981).
12. Calos, M. P. & Miller, J. H. *Cell* **20**, 579–595 (1980).
13. Temin, H. M. *Cell* **21**, 599–600 (1980).
14. Dawid, I. B., Long, E. O., DiNocera, P. P. & Pardue, M. L. *Cell* **25**, 399–408 (1981).
15. Van Arsdell, S. W. *et al. Cell* **26**, 11–17 (1981).
16. Jagadeeswaran, P., Forget, B. G. & Weissman, S. M. *Cell* **26**, 141–142 (1981).
17. Streek, R. E. *Nature* **298**, 767–769 (1982).
18. Kominami, R., Mishima, Y., Urano, Y., Sukai, M. & Muramatsu, M. *Nucleic Acids Res.* **10**, 1963–1979 (1982).
19. Maxam, A. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560–564 (1977).
20. Gross-Bellard, M., Oudet, P. & Chambon, P. *Eur. J. Biochem.* **36**, 32–38 (1973).
21. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
22. Denhardt, D. *Biochem. biophys. Res. Commun.* **23**, 641–652 (1966).

Hypomethylation distinguishes genes of some human cancers from their normal counterparts

Andrew P. Feinberg & Bert Vogelstein

Cell Structure and Function Laboratory, The Oncology Center, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA

It has been suggested that cancer represents an alteration in DNA, heritable by progeny cells, that leads to abnormally regulated expression of normal cellular genes; DNA alterations such as mutations^{1,2}, rearrangements^{3–5} and changes in methylation^{6–8} have been proposed to have such a role. Because of increasing evidence that DNA methylation is important in gene expression (for review see refs 7, 9–11), several investigators have studied DNA methylation in animal tumours, transformed cells and leukaemia cells in culture^{8,12–30}. The results of these studies have varied; depending on the techniques and systems used, an increase^{12–15}, decrease^{20–24}, or no change^{25–29} in the degree of methylation has been reported. To our knowledge, however, primary human tumour tissues have not been used in such studies. We have now examined DNA methylation in human cancer with three considerations in mind: (1) the methylation pattern of specific genes, rather than total levels of methylation, was determined; (2) human cancers and adjacent analogous normal tissues, unconditioned by culture media, were analysed; and (3) the cancers were taken from patients who had received neither radiation nor chemotherapy. In four of five patients studied, representing two histological types of cancer, substantial hypomethylation was found in genes of cancer cells compared with their normal counterparts. This hypomethylation was progressive in a metastasis from one of the patients.

The technique we used is based on the fact that certain restriction endonucleases, which cleave at sites containing the sequence 5'-CG-3', do not function if the internal cytosine residue is methylated^{31–33}. Cytosine is the only significantly modified base in mammalian DNA³⁴, and the two-base sequence 5'-CG-3' can be heritably methylated^{35–38}. Thus, enzymes that discriminate between methylated and unmethylated CG sequences (for example *HpaII* and *HhaI*) are sensitive indicators of methylation^{32,33}. Indeed, these enzymes have been used in many studies to show an inverse correlation between gene expression and methylation (reviewed in refs 7, 9–11), and in other studies to detect methylation differences in animal tumours and transformed cells^{8,18,22–24,29,30}. In the present study, DNA was purified from primary human cancers and from the adjacent normal tissues of the same patients. After cleavage with *HpaII* or *HhaI*, the DNA samples were transferred to nitrocellulose filters by the method of Southern^{39,40}. The level of methylation of specific genes was assessed by hybridization of these filters to ³²P-labelled DNA probes. The probes were made from the cDNA clones of human growth hormone⁴¹, γ -globin⁴² and α -globin⁴², which were chosen on the following bases: (1) the genes to be detected should not be expressed in the normal tissues from which the cancers were derived; (2) the genes should be from widely scattered sites in the genome;

and (3) *HpaII* and *HhaI* cleavage sites should be present in the regions of the genes.

The first cancer studied was a grade D (ref. 43), moderately well differentiated adenocarcinoma of the colon from a 67-year-old male. Tissue was obtained from the cancer itself and also from colonic mucosa stripped from the colon at a site just outside the histologically proven tumour margin. Figure 1 shows the pattern of methylation of the studied genes. Before digestion with restriction enzymes, all DNA samples used in the study had a size >25,000 base pairs (bp). After *HpaII* cleavage, hybridization with a probe made from a cDNA clone of human growth hormone (HGH) showed that significantly more of the DNA was digested to low-molecular weight fragments in DNA from the cancer (labelled C in Fig. 1) than in DNA from the normal colonic mucosa (labelled N). In the hybridization conditions used, the HGH probe detected the human growth hormone genes as well as the related chorionic somatotropin

Table 1 Quantitation of methylation of specific genes in human cancers and adjacent analogous normal tissues

Patient	Carcinoma	Probe	Enzyme	% Hypomethylated fragments		
				N	C	M
1	Colon	HGH	<i>HpaII</i>	<10	35	—
			<i>HhaI</i>	<10	39	—
			<i>HpaII</i>	<10	52	—
		γ -Globin	<i>HhaI</i>	<10	39	—
			<i>HpaII</i>	<10	<10	—
			<i>HhaI</i>	<10	<10	—
2	Colon	HGH	<i>HpaII</i>	<10	76	—
			<i>HhaI</i>	<10	85	—
			<i>HpaII</i>	<10	58	—
		γ -Globin	<i>HhaI</i>	<10	23	—
			<i>HpaII</i>	<10	<10	—
			<i>HhaI</i>	<10	<10	—
3	Colon	HGH	<i>HpaII</i>	<10	41	—
			<i>HhaI</i>	<10	38	—
			<i>HpaII</i>	<10	50	—
		γ -Globin	<i>HhaI</i>	<10	22	—
			<i>HpaII</i>	<10	<10	—
			<i>HhaI</i>	<10	<10	—
4	Colon	HGH	<i>HpaII</i>	<10	<10	—
			<i>HhaI</i>	<10	<10	—
			<i>HpaII</i>	<10	<10	—
		γ -Globin	<i>HhaI</i>	<10	<10	—
			<i>HpaII</i>	<10	<10	—
			<i>HhaI</i>	<10	<10	—
5	Lung	HGH	<i>HpaII</i>	<10	44	86
			<i>HhaI</i>	<10	46	83
			<i>HpaII</i>	<10	82	>90
		γ -Globin	<i>HhaI</i>	<10	<10	<10
			<i>HpaII</i>	24	78	>90
			<i>HhaI</i>	<10	13	23

Autoradiographs such as those shown in Figs 1–3 were scanned with a Clifford Densicom Model 445 densitometer and the scans digitized using a Hewlett-Packard 9874A digitizer. For each enzyme and probe, the detectable fragments were divided into two groups. Group I consisted of the high-molecular weight fragments that were the major fragments detected in the digests of the normal tissues. Group II consisted of more fully digested fragments of lower molecular weight. The proportion II/I+II was determined for each digest and is recorded as % hypomethylated fragments. As controls for the quantitation, varying amounts of undigested DNA or *MspI* digested DNA were pooled, co-electrophoresed, blotted to nitrocellulose and hybridized with each of the probes. There was a nearly linear relationship ($r=0.82$) between the proportion of *MspI*-digested DNA in the sample and the ratio II/I+II over a 10-fold range. Because the film response was linear only over this range, however, ratios between 0% and 10% are listed as <10%, and ratios between 90% and 100% are listed as >90%. N, normal tissue; C, primary cancer; M, metastasis.

Fig. 1 Methylation pattern of adenocarcinoma of the colon and normal colonic mucosa from patient 1. Lanes C and N were prepared from DNA digests of the colon carcinoma and normal colonic mucosa, respectively. *a*, *HpaII* digest, HGH probe; *b*, *HhaI* digest, HGH probe; *c*, *MspI* digest, HGH probe; *d*, *HpaII* digest, γ -globin probe; *e*, *HhaI* digest, γ -globin probe; *f*, *HpaII* digest, α -globin probe; *g*, *HhaI* digest, α -globin probe. The asterisks on the left of *a*–*g* indicate molecular weight markers (*HindIII*-digested bacteriophage λ DNA, of sizes 9,500, 6,700, 4,400, 2,000 and 570 bp from top to bottom, respectively).

Methods: Tissues were frozen in liquid nitrogen, then pulverized and the DNA extracted essentially as described elsewhere⁵¹. The DNA was cleaved with 50-fold excess of restriction endonuclease, as assessed by the amount of restriction endonuclease required to digest pBR322 DNA included in a matched duplicate digest of the human DNA. DNA digests (5 μ g per lane as assessed by a fluorimetric assay⁵²) were electrophoresed on 0.8–1.5% agarose gels at 70 V for 4 h, then transferred to nitrocellulose by the modification of Southern's³⁹ procedure described by Wahl *et al.*⁴⁰. pBR322 plasmids containing cDNA inserts of human growth hormone (chGH800/pBR322; ref. 41), human γ -globin (JW101; ref. 42) and human α -globin (JW101; ref. 42), were grown in L-broth, and plasmid DNA was isolated by standard techniques^{53,54}. The inserts from these plasmids were purified and the DNA labelled with ³²P-dCTP to a specific activity of 10⁹ d.p.m. μ g⁻¹ using a technique described elsewhere⁵⁵. The blots obtained by this method were identical to those obtained when probes were labelled by nick-translation, but exposure times were significantly reduced. The probes were hybridized to the filters for 36–60 h and then washed according to a protocol supplied by K. Peden⁵⁶. The autoradiographs shown were exposed for 2–4 days using pre-exposed Kodak XAR-5 film with DuPont Lightning Plus intensifying screens⁵⁷.

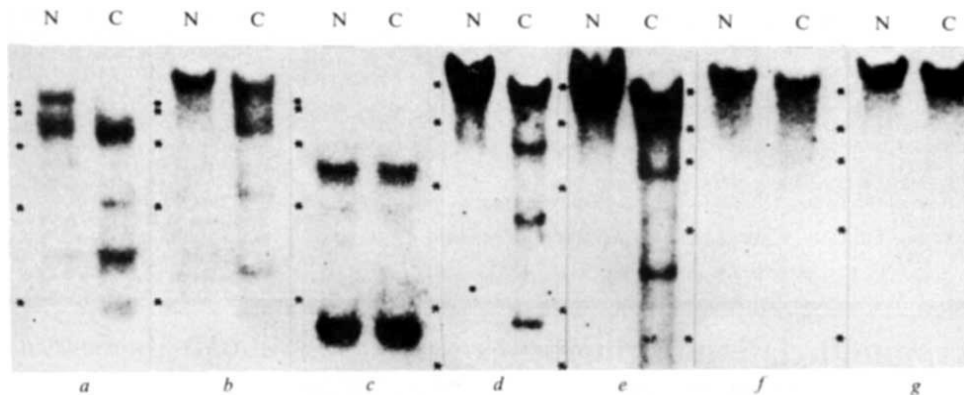


Fig. 2 Methylation patterns of human adenocarcinoma and normal colonic mucosa from patients 2 and 3. Autoradiographs were prepared from blots made as described in Fig. 1 legend. Lanes labelled C and N were prepared from DNA digests of the colon carcinoma and normal colonic mucosa, respectively. *a*, *HpaII*, HGH probe, patient 2; *b*, *HhaI*, HGH probe, patient 2; *c*, *HpaII*, γ -globin probe, patient 2; *d*, *HhaI*, γ -globin probe, patient 2; *e*, *HhaI*, HGH probe, patient 3; *f*, *HpaII*, γ -globin probe, patient 3; *g*, *HhaI*, γ -globin probe, patient 3. Asterisks (molecular weight markers) are as in Fig. 1.

Fig. 3 Methylation pattern of small cell carcinoma of the lung, normal lung, and liver metastasis from patient 5. Autoradiographs were prepared as described in Fig. 1 legend. Lanes labelled C, M and N were prepared from DNA digests of the lung carcinoma, liver metastasis and normal lung, respectively. *a*, *HpaII*, HGH probe; *b*, *MspI*, HGH probe; *c*, *HpaII*, γ -globin probe; *d*, *HpaII*, α -globin probe; *e*, *HhaI*, α -globin probe. Asterisks (molecular weight markers) are as in Fig. 1.

genes⁴⁴. Figure 1*a* shows that the internal cytosine in the *HpaII* recognition sequence 5'-CCGG-3' of the cancer cells was substantially less methylated than that of their normal counterparts in the region of one or more of these genes. Cleavage with *HhaI* showed that the internal cytosine in the sequence 5'-GCGC-3' was also significantly hypomethylated in the cancer cells, as assessed with the HGH probe, compared with the normal counterparts (Fig. 1*b*). In contrast, the HGH probe revealed no difference between the normal and cancer cells in methylation of the external cytosine residues in the sequence 5'-CCGG-3'. This was shown by treating the DNA samples with *MspI*, an enzyme that cleaves the recognition sequence 5'-CCGG-3' regardless of the state of methylation of the internal cytosine⁴⁵, but does not cleave if the external cytosine is methylated^{46,47}. In all the cases examined here (Figs 1*c*, 3*b* and

unpublished data), no differences in the *MspI* patterns between the DNA from the normal and cancerous tissues were detected with either the HGH probe or the globin probes. This result is of interest because it has been shown, at least in an experimental system, that the sequence 5'-CmCGG-3' is stably inherited, while the sequence 5'-mCCGG-3' is not⁴⁸. In addition, for all the patients, no differences between the normal and cancerous tissues were detected with the enzymes *HindIII* or *EcoRI*, which are not sensitive to methylcytosine residues³³.

Differences in methylated sites were also found in the region of the γ -globin gene in patient 1, using the 5'-mCG-3'-sensitive enzymes (Fig. 1*d*, *e*). No substantial differences in methylation were found in the region of the α -globin gene (Fig. 1*f*, *g*).

The second patient studied was a 63-yr-old male with a grade D, poorly differentiated colonic adenocarcinoma. This patient

also showed clear-cut decreases in the degree of methylation of the growth hormone-related genes and the γ -globin gene (Fig. 2a-d). However, the pattern of the bands was in some cases different from that of the first patient (compare Figs 1 and 2). There was a heterogeneous pattern of methylation from cell to cell within each tumour, as indicated by the presence of several low-molecular weight bands of varying intensity on Southern hybridization. In contrast, these sites were almost uniformly methylated in normal colonic mucosa. These results are not consistent with a clonal (unique) methylation pattern in the tumours, but suggest that a varying proportion of the cells in each cancer have failed to perform the normal methylation at individual 5'-CG-3' sites. The resultant heterogeneous pattern of methylation has obvious parallels to the heterogeneity in biological properties often noted in cancer⁴⁹.

Patient 3 was a 75-yr-old female with a grade D, moderately well differentiated colonic adenocarcinoma. Hybridization of both HGH and γ -globin probes to *HpaII* and *HhaI* DNA digests also revealed substantial hypomethylation in the cancer in the region of these genes, compared with normal tissue (Fig. 2e-g). Patient 4, a 69-yr-old female with a grade D, poorly differentiated colonic adenocarcinoma, showed no differences in methylation between the cancerous and normal tissues for the three genes studied (Table 1). This analysis obviously does not preclude DNA hypomethylation in other areas of the genome in this patient.

A second type of cancer, small cell carcinoma of the lung, was removed from patient 5, a 53-yr-old male. DNA was extracted from the primary cancer, from normal lung tissue, and from a liver metastasis. Each of the three probes showed substantial hypomethylation at the *HpaII* and *HhaI* sites in the cancer tissue compared with normal lung tissue (Fig. 3). It is significant that the liver metastasis, which was undoubtedly derived from the same cell type as the primary tumour, showed even greater hypomethylation than the primary cancer. Thus, hypomethylation was progressive from primary cancer to metastasis in this patient. This pattern of increasing degree of hypomethylation, from normal to primary cancer to metastasis, was evident with all the probes used (Fig. 3a, c-e).

The relative degree of hypomethylation of the three genes studied was assessed by densitometry. Four of the five cancers showed 3-9-fold differences in the degree of methylation in at least two of the gene regions studied (compared with the normal tissue of the same patient), as assessed by the method described in Table 1. Methods for detecting general hypomethylation might not have shown differences between the DNA samples from the cancerous and normal tissues studied here. For example, Fig. 4 shows an ethidium bromide-stained gel containing *HpaII*- and *MspI*-digested DNA samples from patient 5. This gel was used to make the blots shown in Fig. 3a, b. No gross differences in methylation were detected among the normal, cancerous and metastatic tissues (Fig. 4), even though three different probes all showed significant hypomethylation on Southern analysis (Fig. 3). Therefore, our data suggest that specific sites in genes of cancer cells may be hypomethylated even in the absence of a general hypomethylation detectable with less sensitive techniques.

Thus, we observed substantial hypomethylation in several specific genomic regions for four of five human cancers, compared with adjacent normal cells from the same patients. A metastasis from one patient showed an even greater degree of hypomethylation than the primary tumour. Further studies are required, using other tumours and different probes, to determine the prevalence of this form of genomic alteration in neoplasia. However, this study clearly shows that such DNA alterations exist in at least some human cancers. These alterations may be widespread, as the genes studied are localized on three different chromosomes⁵⁰. Whether such alterations underlie the abnormal pattern of gene expression found in cancer remains a subject for future investigation.

We thank Drs Stephen B. Baylin and Herbert C. Hoover for help in obtaining the tissues used in this study, and Drs John

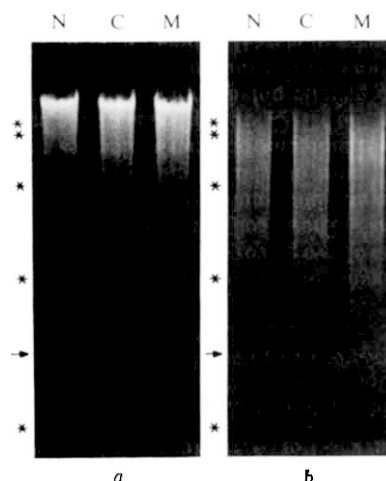


Fig. 4 Ethidium bromide-stained gel of DNA samples digested with *HpaII* and *MspI*. DNAs were prepared from lung carcinoma (C), liver metastasis (M) and normal lung (N) of patient 5, and 5 μ g of each were digested with *HpaII* (a) or *MspI* (b) as described in Fig. 1 legend. The DNA samples were electrophoresed on a 1% agarose gel, stained with ethidium bromide (2 μ g ml⁻¹) and photographed under UV light. The DNA was then transferred to nitrocellulose for the experiments shown in Fig. 3a, b. Asterisks (molecular weight markers) are as in Fig. 1. The arrow indicates the bromophenol blue dye marker.

A. Phillips III, Peter H. Seeburg, Haig H. Kazazian and Corinne Boehm for providing us with cloned DNA fragments. This investigation was supported by USPHS grants 09071 and 31053, awarded by the NCI, DHHS.

Received 27 July; accepted 1 November 1982.

- McCann, J. & Ames, B. N. *Proc. natn. Acad. Sci. U.S.A.* **73**, 950-954 (1976).
- Meselson, M. & Russel, K. in *Origins of Human Cancer* Vol. 4 (eds Hiatt, H. H., Watson, J. D. & Winsten, J. A.) 1473-1481 (Cold Spring Harbor, New York, 1977).
- Sager, R. *Nature* **282**, 447-448 (1979).
- Cairns, J. *Nature* **289**, 353-357 (1981).
- Humphreys, P. *Nature* **293**, 146-148 (1981).
- Holliday, R. *Br. J. Cancer* **40**, 513-521 (1979).
- Ehrlich, M. & Wang, R. Y.-H. *Science* **212**, 1350-1357 (1981).
- Sager, R. in *Tumor Cell Heterogeneity: Origins and Implications* (eds Owens, A. H., Coffey, D. S. & Baylin, S. B.) 411-423 (Academic, New York, 1982).
- Razin, A. & Riggs, A. D. *Science* **210**, 604-610 (1980).
- Doerfler, W. *J. gen. Virol.* **57**, 1-20 (1981).
- Felsenfeld, G. & McGhee, J. *Nature* **296**, 602-603 (1981).
- Silber, R. *et al. Biochim. biophys. Acta* **123**, 638-640 (1966).
- Desai, L. S., Wulff, U. C. & Foley, G. E. *Expl Cell Res.* **65**, 260-263 (1971).
- Shirakawa, S. & Saunders, G. F. *Proc. Soc. exp. Biol. Med.* **138**, 369-372 (1971).
- Rubery, E. D. & Newton, A. A. *Biochim. biophys. Acta* **324**, 24-36 (1973).
- Gunther, U., Schweiger, M., Stupp, M. & Doerfler, W. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3923-3927 (1976).
- Fedorov, N. A., Kuzmichev, V. A., Kritsky, G. A. & Vinogradova, Yu. E. *Biokhimiya* **42**, 1020-1023 (1977).
- Nakhasi, H. L., Lynch, K. R., Dolan, K. P., Unterman, R. D. & Feigelson, P. *Proc. natn. Acad. Sci. U.S.A.* **78**, 834-837 (1981).
- Nass, M. M. K. *J. molec. Biol.* **80**, 155-175 (1973).
- Burtseva, N. N., Axizov, Yu. M., Itkin, B. Z. & Vanyushin, B. F. *Biokhimiya* **42**, 1690-1696 (1977).
- Lapeyre, J.-N. & Becker, F. F. *Biochem. biophys. Res. Commun.* **87**, 698-705 (1979).
- Cohen, J. C. *Cell* **19**, 653-662 (1980).
- Reilly, J. G., Thomas, C. A. Jr & Sen, A. *Biochim. biophys. Acta* **697**, 53-59 (1982).
- Nakhasi, H. L. *et al. J. biol. Chem.* **257**, 2726-2729 (1982).
- Gantt, R., de Oca, F. M. & Evans, V. J. *In Vitro* **8**, 288-294 (1973).
- Singer, J., Stellwagen, R. H., Roberts-Ems, J. & Riggs, A. D. *J. biol. Chem.* **252**, 5509-5513, 1977.
- Browne, M. J. & Burdon, R. H. *Nucleic Acids Res.* **4**, 1025-1037 (1977).
- Diala, E. S. & Hoffman, R. M. *Biochem. biophys. Res. Commun.* **104**, 1489-1494 (1982).
- McKeon, C., Ohkubo, H., Pastan, I. & de Crombrughe, B. *Cell* **29**, 203-210 (1982).
- Kuhlmann, I. & Doerfler, W. *Virology* **118**, 169-180 (1982).
- Mann, M. B. & Smith, H. O. *Nucleic Acids Res.* **4**, 4211-4221 (1977).
- Bird, A. P. & Southern, E. M. *J. molec. Biol.* **118**, 27-47 (1978).
- van der Ploeg, L. H. T. & Flavell, R. A. *Cell* **19**, 947-958 (1980).
- Hall, R. H. *The Modified Nucleosides in Nucleic Acids*, 281-294 (Columbia, New York, 1971).
- Bird, A. P. *J. molec. Biol.* **118**, 49-60 (1978).
- Pollack, Y., Stein, R., Razin, A. & Cedar, H. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6463-6467 (1980).
- Wigler, M., Levy, D. & Peruchio, M. *Cell* **24**, 33-40 (1981).
- Harland, R. M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2323-2327 (1982).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
- Wahl, G. M., Stern, M. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3683-3687 (1979).
- Martial, J. A., Halliwell, R. A., Baxter, J. D. & Goodman, H. M. *Science* **205**, 602-607 (1979).
- Wilson, J. T. *et al. Nucleic Acids Res.* **5**, 563-581 (1978).

43. Sugarbaker, P. H., Macdonald, J. S. & Gunderson, L. L. in *Cancer: Principles and Practice of Oncology* (eds DeVita, V. T. Jr, Hellman, S. & Rosenberg, S. A.) 643-723 (Lippincott, Philadelphia, 1982).
44. Moore, D. D., Conkling, M. A. & Goodman, H. M. *Cell* **29**, 285-286 (1982).
45. Waalwijk, C. & Flavell, R. A. *Nucleic Acids Res.* **5**, 3231-3236 (1978).
46. Sneider, T. W. *Nucleic Acids Res.* **8**, 3829-3840 (1980).
47. van der Ploeg, L. H. T., Groffen, J. & Flavell, R. A. *Nucleic Acids Res.* **8**, 4563-4574 (1980).
48. Stein, R., Gruenbaum, Y., Pollack, Y., Razin, A. & Cedar, H. *Proc. natn. Acad. Sci. U.S.A.* **79**, 61-65 (1982).
49. Owens, A. H., Coffey, D. S. & Baylin, S. B. *Tumor Cell Heterogeneity: Origins and Implications* (Academic, New York, 1982).
50. McKusick, V. A. *Cytogenet. Cell Genet.* **32**, 1-343 (1982).
51. Gross-Bellard, M., Oudet, P. & Chambon, P. *Eur. J. Biochem.* **36**, 32-38 (1973).
52. Brunk, C. F., Jones, K. C. & James, T. W. *Analyt. Biochem.* **92**, 497-500 (1979).
53. Colman, A., Byers, M. J., Primrose, S. B. & Lyons, A. *Eur. J. Biochem.* **91**, 303-310 (1978).
54. Birnboim, H. C. & Dolly, J. *Nucleic Acids Res.* **7**, 1513-1523 (1979).
55. Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* (submitted).
56. Peden, K., Mounts, P. & Hayward, G. S. *Cell* **31**, 71-80 (1982).
57. Laskey, R. A. & Mills, A. D. *FEBS Lett.* **82**, 314-316 (1977).

Chloroplast DNA exists in two orientations

Jeffrey D. Palmer

Carnegie Institution of Washington, Department of Plant Biology, Stanford, California 94305, USA

An almost universal feature of the circular chloroplast genome is a large inverted repeat sequence, some 10-25 kilobase pairs (kb) in size, which separates the remainder of the molecule into single copy regions of ~80 kb and 20 kb¹⁻³. A number of physical properties—formation of head-to-head dimers¹, copy-correction between the inverted repeat segments^{1,3,4}, resistance to intramolecular recombinational loss¹⁻³, and maintenance of a highly stable chloroplast genome resistant to rearrangement²—have been attributed to the presence of this large inverted repeat. However, one property which an inverted repeat might be expected to confer—reversal of polarity of the single copy sequences located between the repeats¹—has not yet been demonstrated for the chloroplast genome. I now show that chloroplast DNA prepared from a single plant of common bean (*Phaseolus vulgaris*) consists of two equimolar populations of molecules differing only in the relative orientation of their single copy sequences. A model is presented to explain these results, and comparisons are made to similar cases of inversion heterogeneity in 2-micrometre plasmid DNA from yeast^{5,6} and in herpes simplex virus DNA^{7,8}.

The polarity of the unique sequences between the inverted repeats found in chloroplast DNA molecules can be determined by restriction fragment analysis, but it requires the use of an enzyme that does not cleave within the inverted repeat⁹ but does cleave the unique regions asymmetrically with respect to the inverted repeat. Recently, restriction enzyme maps have been constructed for the chloroplast genome of the common bean (*P. vulgaris*), and it has been found that *SalI* does not cleave the inverted repeat, but does cut in an asymmetric fashion four times in the large and once in the small single copy regions (Fig. 1).

Digestion of common bean chloroplast DNA with *SalI* yields seven fragments, the sizes of which add up to 249 kb (Fig. 2, track 2). The four largest *SalI* bands appear with approximately half the intensity of the three smallest bands. This same fragment pattern was seen in multiple digests of the same DNA using an excess of *SalI*, and also in complete *SalI* digests of chloroplast DNA prepared from a different cultivar (Hawkesbury Wonder) from that shown in Fig. 1 (Kentucky Wonder Bean). However, analysis with a number of other restriction enzymes, including *PstI* (Fig. 1) and *XhoI*^{10,11}, gives a genome size of 150 kb, ~100 kb lower than that obtained with *SalI*. The simplest explanation for this discrepancy is that molecular heterogeneity is revealed by *SalI*, which does not cleave the inverted repeat (Fig. 1), but not by enzymes such as *PstI* (Fig. 1) and *XhoI*^{10,11}, which do.

This was confirmed by hybridizing cloned mung bean *PstI* fragments of 18.8 kb and 9.7 kb, which are defined by the same

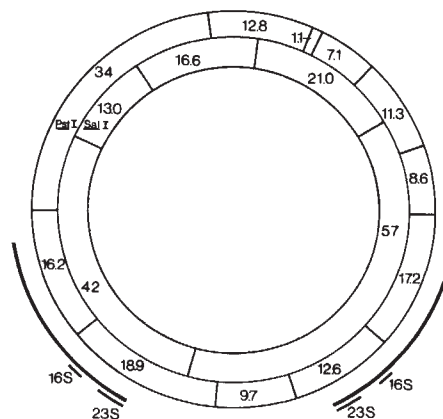


Fig. 1 Restriction map of common bean chloroplast DNA¹⁰ showing the position of the inverted repeat (long heavy lines) and a single orientation of the unique sequence regions. Fragment sizes are in kb.

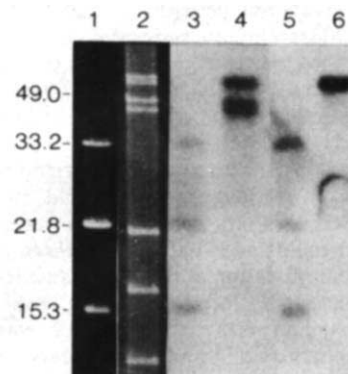


Fig. 2 Physical heterogeneity of common bean chloroplast DNA. Chloroplast DNA was purified²³ from a single common bean plant (*Phaseolus vulgaris* cv. Kentucky Wonder Bean). *EcoRI* and *SalI* restriction fragments of phage λ DNA, together with intact λ DNA (track 1), and *SalI* fragments of common bean chloroplast DNA (track 2) were separated by electrophoresis in a 0.4% agarose gel²³. Duplicate nitrocellulose filter blots of the gel were prepared according to a modification²³ of the bidirectional transfer method of Smith and Summers²⁴. Filters were hybridized²³ using as probes ³²P-nick-translated²⁵ cloned²⁶ mung bean chloroplast DNA *PstI* fragments of 18.8 kb (tracks 3 and 4) and 9.7 kb (tracks 5 and 6). Size of λ fragments are given in kb. Hybridization of the pBR322-chloroplast DNA recombinant plasmids to λ phage restriction fragments in tracks 3 and 5 results from a small amount of cross-homology between pBR322 and λ DNAs²⁷.

PstI sites as the common bean fragments of 18.9 kb and 9.7 kb¹⁰ (Fig. 1), to nitrocellulose filter blots of the gel lanes shown in tracks 1 and 2 of Fig. 2. The 18.8-kb fragment hybridizes to the 42-kb and 46-kb *SalI* fragments, and also to an unresolved band which probably represents the two largest *SalI* fragments of 57 and 53 kb (Fig. 2, track 4). The 9.7-kb probe hybridizes only to the broad 57-53-kb band (Fig. 2, track 6). The most reasonable explanation for these data is that common bean chloroplast DNA consists of two distinct molecular species, present in approximately equal proportions, which differ only with respect to the relative orientation of their single-copy DNA regions (Fig. 3, bottom).

The top portion of Fig. 3 presents a schematic representation of the hypothetical intermediates in the interconversion of the two circular chloroplast DNA isomers. It is envisioned that reciprocal recombination between the paired inverted repeat segments leads to direct interconversion of the dumbbell-shaped intermediates, which may then relax to give the circular conformation typically seen in the electron microscope^{1,12,13}. That no dumbbell-shaped molecules have been observed in the electron microscope^{1,12,13} suggests either that such structures are fragile and difficult to isolate or that the vast majority of chloroplast DNA molecules are in the unpaired, circular confor-

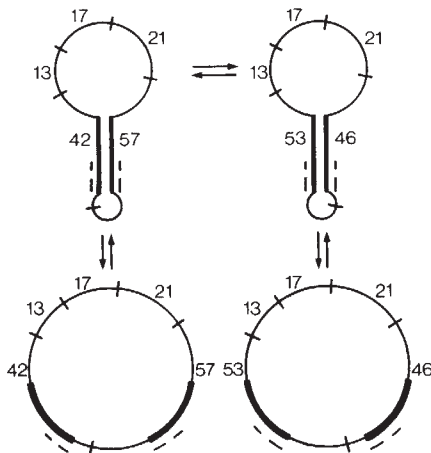


Fig. 3 Model for chloroplast DNA isomerization via intra-molecular recombination between segments of the inverted repeat. *SalI* sites and fragments (in kb), as well as locations of the inverted repeat (long, heavy lines) and ribosomal RNA genes (short lines at the small single-copy end of the inverted repeat) are given for common bean chloroplast DNA (see Fig. 1).

mation. Whatever the precise mechanism, interconversion must occur frequently enough on a developmental time scale to produce equal amounts of the two chloroplast DNA isomers in a single plant. Alternatively, interconversion may be quite rare, or even absent, and instead some physical mechanism may ensure a regular and equal distribution of the two isomers to each daughter plastid at the time of plastid division. However, this seems unlikely given the rapid sorting out both of differentially marked chloroplast genomes, in the green alga *Chlamydomonas reinhardtii*, and also of differentially marked chloroplasts, in somatic cell fusions of flowering plants^{3,14,15}.

It seems unlikely that common bean chloroplast DNA is atypical of inverted repeat-containing chloroplast DNAs, and it is expected, given the appropriate diagnostic restriction enzymes, that other inverted repeat-containing chloroplast genomes will be shown to exist in two orientations. Indeed, we have recently found that chloroplast DNA from each of three species in the fern genus *Osmunda*, whose inverted repeats are no more than half the size of that found in a typical angiosperm¹⁶, also exists as two equimolar populations differing only in the relative orientation of their single copy sequences (unpublished data).

The same type of orientation heterogeneity observed for chloroplast DNA has been seen in other systems. Most notably, most strains of yeast in the genus *Saccharomyces* possess a 6.3-kb plasmid DNA (2- μ m circle) whose 0.6-kb inverted repeat sequences also mediate frequent single-copy flip-flops⁵. Broach *et al.*⁶ have shown that high frequency interconversion in the 2- μ m circle is mediated by a specialized site-specific recombination system encoded by a specific gene (FLP) carried on the plasmid itself. A similar situation exists in the linear DNA of herpes simplex virus and related viruses, where site-specific recombination between the two sets of inverted repeats that bracket large and small single-copy sequences leads to the formation of four equimolar populations of isomerically inverted molecules^{7,8,17}.

These comparisons raise important questions concerning the site of chloroplast DNA recombination and isomerization, and the nature and genetic locus of the recombination system involved. If, by analogy to the yeast system, the chloroplast genome does encode its own specialized recombination system, what is the fate of that system in those legumes such as pea^{1,18,19} that have deleted one segment of the inverted repeat? Has the system itself been deleted? Is it still present but no longer functional and are the genes now evolving as pseudogenes? Or is the system still operating, perhaps in an altered manner somehow responsible for the greatly increased incidence of major structural rearrangements observed in the pea genome^{2,19}?

Is the existence of two populations of chloroplast DNA molecules differing only in their single copy orientations of any functional importance to the plastid and the plant? Two possible roles have been suggested for the switching system in the yeast 2- μ m plasmid⁵. Interconversion may assist in propagation of the plasmid, by helping to resolve catenated molecules that arise during replication, or alternatively, may act as a genetic switch, allowing alternate transcription of flanking genes in a manner analogous to regulatory switching in the *H-2* locus of *Salmonella*²⁰ and the G loop of phage Mu²¹. Analysis of the transcripts produced at the chloroplast DNA single copy-inverted repeat junctions may provide a clue regarding this latter possibility, although the observation²² of a 1-kb deletion which spans one of the large single copy-inverted repeat junctions in *Atriplex halimus* argues against this function at least one such junction.

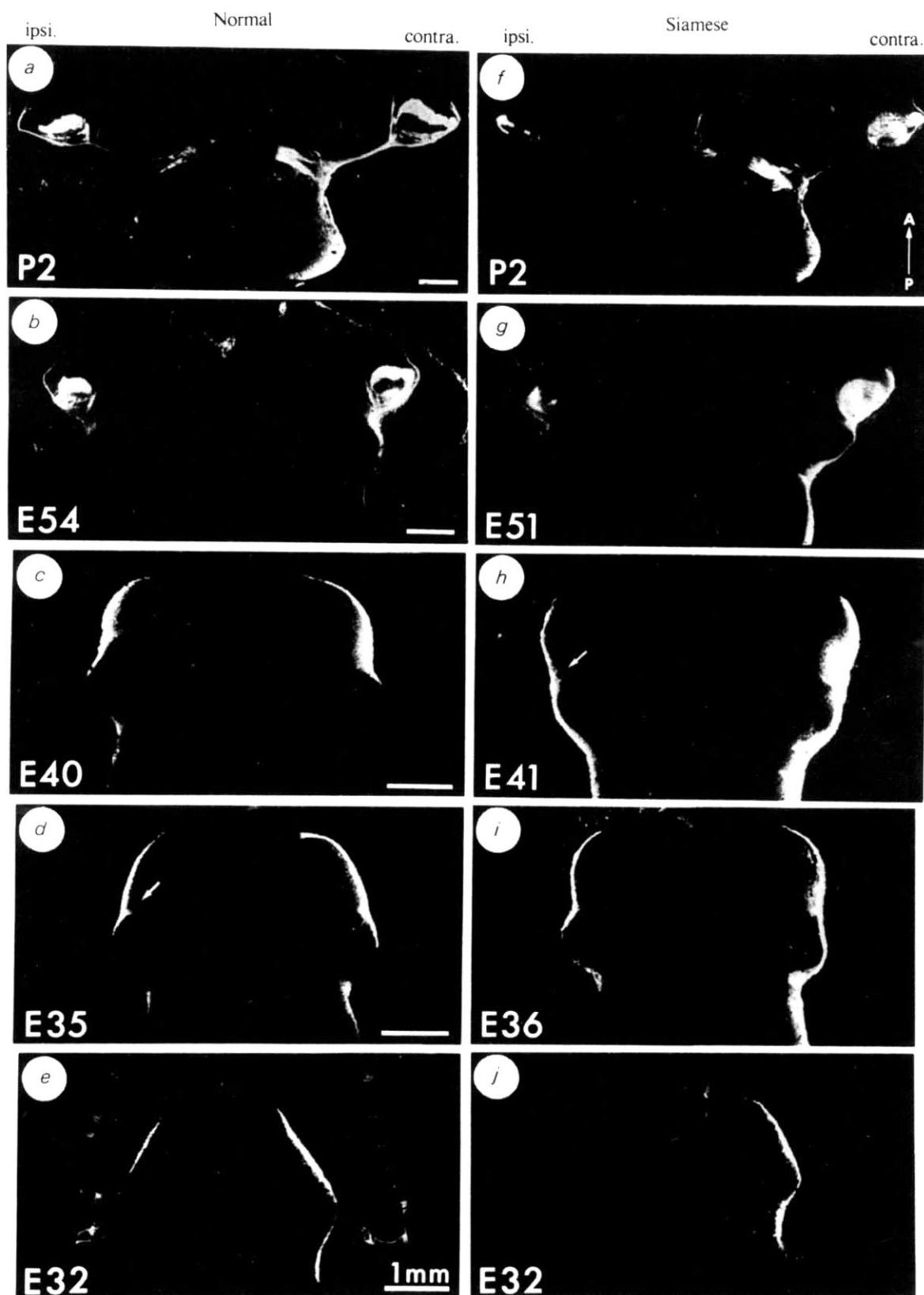
The interconversion phenomenon documented here may be added to a growing list of physical properties conferred on the chloroplast genome by its distinctive large inverted repeat. Formation of head-to-head circular dimers, a feature unique to inverted repeat-containing chloroplast genomes¹, and polarity reversals of single copy sequences appear to be relatively frequent events that arise from recombination within the inverted repeats (intermolecularly in the former case and intramolecularly in the latter) and which do not have any known evolutionary significance. In contrast, the concerted evolution^{1,3,4} of the two segments of the inverted repeat is of obvious evolutionary import. What is unclear is whether this process is dependent on the physical structure of the inverted repeat and the intramolecular recombinational and conversion opportunities it affords, or whether it results simply from stochastic matching and conversion properties which operate on the entire genome. That the inverted repeat is such a highly conserved structure, with only a single loss observed among plants representing several hundred million years of evolution^{2,3}, may be viewed as an allied property to polarity reversal. That is, intramolecular recombination between inverted repeat segments can at most reverse the orientation of single copy sequences, but cannot delete one of the repeats, as is possible when they are present as direct repeats¹. Finally, the inverted repeat may stabilize the entire chloroplast genome against major sequence rearrangements, so that these events are extremely rare in those genomes that retain the inverted repeat and increase markedly in frequency when the repeat is lost².

I thank William F. Thompson, in whose laboratory this research was done, B. Osorio and H. Edwards for technical assistance and M. Zolan and J. Watson for helpful discussion. This is CIW-DPB publication 792.

Received 15 September; accepted 1 November 1982.

- Kolodner, R. & Tewari, K. K. *Proc. natn. Acad. Sci. U.S.A.* **76**, 41–45 (1979).
- Palmer, J. D. & Thompson, W. F. *Cell* **29**, 537–550 (1982).
- Gillham, N. W., Boynton, J. E. & Harris, E. H. in *DNA Evolution: Natural Selection and Genome Size* (ed. Cavalier-Smith) (Wiley, New York, in the press).
- Gordon, K. H. J., Crouse, E. J., Bohnert, H. J. & Herrmann, R. G. *Theor. appl. Genet.* **61**, 373–384 (1982).
- Broach, J. R. *Cell* **28**, 203–204 (1982).
- Broach, J. R., Guarascio, V. R. & Jayaram, M. *Cell* **29**, 227–234 (1982).
- Roizman, B. *Cell* **16**, 481–494 (1979).
- Mocarski, E. S. & Roizman, B. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7047–7051 (1981).
- Gordon, K. H. J., Crouse, E. J., Bohnert, H. J. & Herrmann, R. G. *Theor. appl. Genet.* **59**, 281–296 (1981).
- Palmer, J. D., Singh, G. P. & Pillay, D. T. N. *Molec. gen. Genet.* (in the press).
- Mumbilla, M., Gordon, K. H. J., Crouse, E. J., Burkard, G. & Weil, J. H. *Gene* (submitted).
- Kolodner, R. & Tewari, K. K. *J. biol. Chem.* **250**, 8840–8847 (1975).
- Koller, B. & Delius, H. *Molec. gen. Genet.* **178**, 261–269 (1980).
- VanWinkle-Swift, K. P. *Curr. Genet.* **1**, 113–125 (1980).
- Scowcroft, W. R. & Larkin, P. J. *Theor. appl. Genet.* **60**, 179–184 (1981).
- Palmer, J. D. & Stein, D. B. *Curr. Genet.* **5**, 165–170 (1982).
- Weststrate, M. W., Geelen, J. L. M. C. & Van der Noordaa, J. *J. gen. Virol.* **49**, 1–21 (1980).
- Chu, N. M., Oishi, K. K. & Tewari, K. K. *Plasmid* **6**, 279–292 (1981).
- Palmer, J. D. & Thompson, W. F. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5533–5537 (1981).
- Zeig, J. & Simon, M. I. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4196–4200 (1980).
- Bukhari, A. & Ambrosio, L. *Nature* **271**, 575–577 (1978).
- Palmer, J. D. thesis, Stanford Univ. (1981).
- Palmer, J. D. *Nucleic Acids Res.* **10**, 1593–1605 (1982).
- Smith, G. E. & Summers, M. D. *Analyt. Biochem.* **109**, 123–129 (1980).
- Maniatis, T., Jeffrey, A. & Kleid, D. G. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1184–1188 (1975).
- Palmer, J. D. & Thompson, W. F. *Gene* **15**, 21–26 (1981).
- Pogue-Geile, K. L., Dassarma, S., King, S. R. & Jaskunas, S. R. *J. Bact.* **142**, 992–1003 (1980).

Erratum



In the letter 'Prenatal misrouting of the retinogeniculate pathway in Siamese cats' by C. J. Shatz and M. Kliot, *Nature* **300**, 525–529 (1982), Figure 2 was incorrectly reproduced: the correct version is given above.

A saga of the Thames

Eric Ashby

In 1778 Joseph Bramah took out a patent for a device which has been the cause of death for multitudes of people. It was not a weapon of war: it was the water closet. By the early nineteenth century it became the aspiration, and then the expectation, of householders in towns to have running water and a water closet.

In their enthusiasm they overlooked a simple and grave consequence. Water companies — there were seven of them in London in the year 1800 — brought water into some 100,000 homes, but there were no arrangements for taking the water out again. In 1800 it was forbidden to discharge domestic waste into the London sewers; their purpose was to collect run-off from the streets. Cesspools became flooded, soil workers could no longer empty them, there was a persistent stench round the most respectable homes. So in 1815 permission was given for the wastes to be discharged into sewers and thence into the Thames. These wastes, together with phenols and ammonia from the Gas Light and Coke Company, deprived the river of oxygen and inoculated it with cholera. But water companies continued to draw drinking water from the Thames. Outbreaks of cholera and typhoid killed tens of thousands of Londoners and (an event politically more effective) MPs were sickened by a dreadful stink from the river in 1858 as it flowed past the Houses of Parliament.

The familiar chain-reaction began: a Metropolitan Commission of Sewers, renewed six times between 1847 and 1864; Bills debated in Parliament; and eventually two huge sewers, one north and one south of the river, to export the pollution into the estuary, to the jurisdiction of Kent and Essex.

This is but one episode in the long and — for a sanitary engineer, exciting — history of the Thames. The river deteriorated again and became a disgrace by 1950. More committees, more recommendations, more wrangling between authorities as to who was to pay and who was to have authority; and a successful solution to the problem, culminating in press reports of salmon caught in 1974 and, in 1978, 53 species of fish recorded at the intake at West Thurrock power station.

All this, in a vague way, is common knowledge. But only in a vague way. The detailed story of the technological problems, the biological research, the conflicts of interest between those who wanted to be able to drink Thames water and those who wanted to discharge wastes

The Restoration of the Tidal Thames. By Leslie B. Wood. Pp.202. ISBN 0-85274-447-1. (Adam Hilger, Bristol/Heyden: 1982.) £22.50, \$49.50.

into it, the administrative complexities in the management of a river exploited at one time by 182 separate sewage, gas and petrochemical works, distilleries and paper mills — this story of the Thames has not yet been fully told. But Leslie Wood has given us, in this brief book, the best introduction to the saga that has yet been written. The book is packed with facts, tables, graphs, plans of sewage works and photographs of some of the men whose expertise and persistence have made the restoration of the Thames one of the great success stories in the politics of the environment.

The author begins with a chapter on the history of the river before 1800, including the fascinating example of what I suppose (to be topical) might be called "punctuated social evolution": the consolidating legislation known as the Bill of Sewers, passed in 1531, and which remained on the statute book, unaltered for 300 years, surviving the industrial revolution. He then takes us through the deterioration of the river from 1800 to 1850, quoting in full Faraday's famous letter to *The Times*, recording his observation that bits of paper dropped into the Thames were invisible "before they had sunk an inch below the surface". He quotes, too (though one

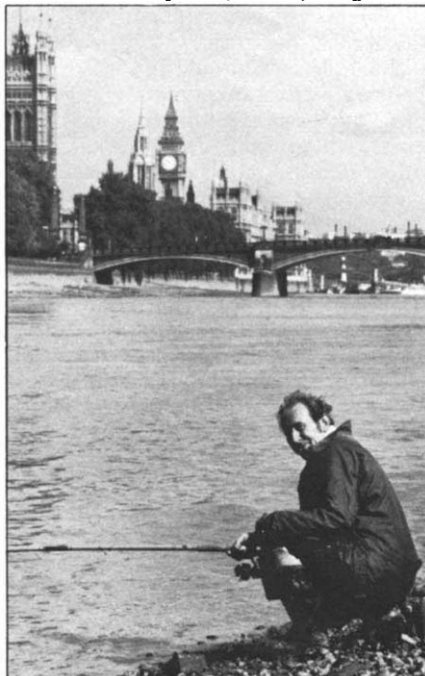
wishes it were in more detail), John Snow's critical observation, with a "control": a record of 71 deaths per 10,000 inhabitants in Southwark, which drew its drinking water from the tidal Thames and, in the same cholera epidemic, only 5 deaths per 10,000 in Lambeth, which drew its drinking water from the Thames too, but above Teddington Weir.

Leslie Wood was an officer employed by the authorities responsible for the management of the tidal Thames and his book is, in a way, a token of appreciation of his colleagues' work. So his main interest is, naturally, not so much in the causes of deterioration but in the measures for restoration. He writes with first-hand knowledge of the strategy being pursued by the Thames Water Authority and he has important things to say — not only for the management of the Thames but about policy for all rivers that run through industrial cities — on the controversial issue of pollution budgets.

Back in 1836 a select committee of Parliament considered a plan for "Rescuing the River Thames from Every Species of Pollution". The managers of the Thames have more modest aims now. Their job is to reconcile many uses of the river, not just to restore it to some Arcadian vision of pristine purity (which in any case it did not have even in the fourteenth century). So instead of "rescuing" the Thames from "Every Species of Pollution" the authority determines what levels of pollution will on one hand enable industry to flourish on the river bank, and on the other hand ensure that the pollutants will be so diluted that they will not harm fish or create a nuisance in any other way. Of particular interest, therefore, is a table on p.169 of the book, setting out consent conditions for the discharge of trade effluents, including such emotive substances as cyanide, chlorine, formaldehyde, cadmium and lead. How wise, not to pretend that these substances could be totally eliminated, but to set conditions for their release which can reasonably be complied with.

To the question: "How clean is a river?" the scientist can give an answer. But the question: "Is the river clean *enough*?" cannot be answered by science alone. It is a political question. Leslie Wood's book is a timely reminder that environmental policy is all about the word "enough". □

Lord Ashby is a Fellow of Clare College, Cambridge. He is joint author with Mary Anderson of *The Politics of Clean Air*, published by the Clarendon Press in 1981.



"How clean is a river?"

'Subtle is the Lord...'

The Science and Life of Albert Einstein Abraham Pais

'Abraham Pais is a distinguished physicist in his own right; his scientific biography of Einstein has clearly been a labour of love. The level of the scientific presentation is high, but so are the rewards... Pais's examination of Einstein's scientific oeuvre will be an important acquisition for both physicists and historians and philosophers of science.' *Nature*.

'A book which is more than the sum of its disparate parts.' *The Observer*.

'This is an extraordinary biography of an extraordinary man.' *Christian Science Monitor*.

'Pais has written a superb history of Einstein's work. It is authoritative, excitingly written, and, above all, suffused with an understanding of the processes and content of twentieth-century research in physics. Anyone interested in the history of physics will have to read this book.' *Steven Weinberg*.

Illustrated £15

Oxford University Press

Circle No.07 on Reader Service Card.

**nature
is available in
Microform.**



**University Microfilms
International**

Please send additional information

for _____

Name _____

Institution _____

Street _____

City _____

State _____

Zip _____

300 North Zeeb Road
Dept. P.R.
Ann Arbor, Mi. 48106

Circle No.56 on Reader Service Card.

An encyclopaedia of natural life

Alwyne Wheeler & James H. Price

Synopsis and Classifications of Living Organisms, 2 vols. Edited by Sybil P. Parker. Vol.1, pp.1,166; Vol.2, pp.1,232. ISBN 0-07-079031-0. (McGraw-Hill: 1982.) £150, \$205.

THESE immense and awesome volumes are certainly, even at first sight, an impressive piece of work, embracing the efforts of 174 authors in covering the entire spectrum of living organisms. Aside from the authors, some 14 section-organizing consultants and 10 editorial staff are listed; quite how many others have been involved is anyone's guess. The galaxy of talents so involved embraces many world authorities on various groups, despite the (not surprising) strong, but by no means total, bias towards the North American continent.

The general format is of chapters dealing with particular groups of organisms, written by individual authors; in complex groups where fields are studied separately, several authors have been called on for different parts of chapters. References for additional reading appear at the foot of each section or chapter, and groups of photographic plates illustrate taxa and organisms as considered appropriate by the compilers. Generally, both references and plates are thoughtfully organized, although some of the non-classic references listed for additional reading are rather dated. Some section references tend to reveal author bias, while other sections have insufficient referencing; in some cases the only citation may be given at class or subclass level and, as a result, only the most general works are cited. An opportunity to provide a lead into the more specialized literature is thus lost.

Some plates do not add a great deal to the textual matter, although any generalist using the books as a primary source of reference will probably be glad of the visual impression they provide. The selection of photographs is also uneven as many of the animals at least are clearly very dead (some of the common insects, like the mole cricket, being pinned specimens), and the omission of some really photogenic groups, such as diatoms, seems surprising. On the other hand this must be the only encyclopaedia of natural life that has no photographs of birds or mammals; in view of the prolific illustrations of these animals already available, this was a justifiable decision.

A profound disadvantage of the unavoidable division of the work into two volumes is two-fold: first, all animal groups do not occur in one volume; second, the first volume does not contain an index. The index provided at the end of Vol.II is very good in content and coverage, but it does mean that its 100 or so

pages (35,000 entries, according to the preface) are not accessible to anyone spatially separated from the second volume. Since both plates and credits for plates are listed in full in front of Vol.1 but under subheads of the two volumes separately, why could not the same have been done for the index, either repeating the whole in both volumes, or as appropriate to the individual volume?

The layout is good, and information is easily absorbed from this type-face once data are located. The shortcoming in inadequate citation of literature aside, these volumes constitute an astonishing achievement and a most useful reference work. Specialists will probably disdain to use it for their own group, but its value in the reference library and to generalists in biology will be immense. □

Alwyne Wheeler and James H. Price are taxonomists on the staff of the British Museum (Natural History), London, with particular interests in fishes and cryptogamic botany respectively.

Two views of nerve cells

P.F. Baker

Physiology and Electrochemistry of Nerve Fibers. By Ichiji Tasaki. Pp.348. ISBN 0-12-683780-5. (Academic: 1982.) \$36, £23.80.

THE author of *Physiology and Electrochemistry of Nerve Fibers*, Dr Ichiji Tasaki, is an experimentalist of the very first rank. It follows that when such an individual chooses to write on a topic to which he has devoted a lifetime of effort, other students of the field will have a high expectation of the result.

It has to be said that these expectations are largely unfulfilled. Indeed, the overwhelming impression is that, with the exception of optical studies of nerve, the author has made no real attempt to come to grips with developments in the rest of the world over the past 15 years. There is, for instance, no mention of the study of single-channel conductances in either nerve cells or artificial membranes. As a result, the book provides only a very one-sided account. Obviously Dr Tasaki has every right to take up a stance that differs from the prevailing dogma; but he owes it to himself to make every effort to interpret his own interesting findings within the accepted framework before rejecting the whole fabric of modern cellular electrophysiology for a rather ill-defined alternative.

Dr Tasaki seems convinced that cellular

neurophysiology has been led astray by too much emphasis being laid upon equivalent circuits and too little consideration of the underlying physical chemistry. This may be a valid criticism of the work of some biophysicists, but not of the field in general where in laboratories throughout the world the analysis of single-channel conductances by the patch clamp technique is being vigorously pursued alongside the search for the molecules underlying these unitary events.

Dr Tasaki's insistence on chemistry is undoubtedly correct, and the growing

understanding of the biochemistry and pharmacology of these putative channel proteins and the realization that their properties depend amongst other things on their ionic environment should be music to his ears. Had he attempted to integrate this new wave of channel chemistry with his own experimental results on the electrical and optical properties of perfused axons, something very special might have emerged. As it is, the book is likely to find only a very limited audience. □

P.F. Baker is Halliburton Professor of Physiology at King's College, University of London.

Darwin and Dohrn

J.Z. Young

Charles Darwin, Anton Dohrn: Correspondence. Edited by Christiane Groeben. Pp.118. (Gaetano Macchiaroli Editore, via Carducci 59, 80121 Naples: 1982.) L7,000.

ANTON Dohrn and Charles Darwin were very different characters and it may come as a surprise to learn that they exchanged many letters in the latter years of Darwin's life. The correspondence began in 1867, when Dohrn sent to Darwin a copy of his paper comparing limbs of crustaceans and insects. Darwin answered kindly, although the views expressed must have seemed rather strange. Dohrn was overjoyed and wrote back a typically enthusiastic letter full of "Compliments and deep Veneration which flows out of the heart of a young ardent champion for truth and Liberty".

Then, in 1869, Dohrn explained his "idea of founding not only Aquariums, but Zoological Stations or Laboratories at different points of our European coast". Darwin replied with typical modesty that "your scheme has my good wishes, but I am sure that you estimated my opinion and judgement much too highly". He also suggested "one caution, as Demosthenes said 'action, action, action' was the soul of eloquence, so is caution almost the soul of science".

And so the correspondence went on, even until the last year of Darwin's life. The letters are most revealing of the character and activities of both men. Dohrn writes with great zest though (excusably) often imperfect English grammar. His topics were partly scientific in the early letters but later mainly concerned with the worries of the Zoological Station. From the start until today it has had financial difficulties and battles with both central and local government. Those who have had letters from Anton Dohrn's son Reinhardt and grandson Peter will recognize that they inherited not only Anton's problems but also the character that overcame them as well as Anton's flair for writing with verve and enthusiasm.

The family and the Station they directed have always had strong international connections. Darwin obviously approved of this and was pleased to give practical support. In 1874 he sent a "subscription of £100, and one of £10 each from my two sons George and Francis". Then, in 1880, near the end of his life, Darwin was awarded the Bressa Prize of L12,000 (£100) and offered it to the Station. In accepting this gift Dohrn decided to devote it to encourage "English Naturalists to come more frequently to Naples". While agreeing to this, Darwin characteristically warned that the sum was so small "it appears to me that it would be prudent in you not to speak of the use to which you propose to put it until you are assured of receiving considerable additions".

Biological rhythms out of harmony

J.T. Enright

Biological Timekeeping. Edited by John Brady. Pp.197. Hbk ISBN 0-521-23307-0; pbk ISBN 0-521-29899-7. (Cambridge University Press: 1982.) Hbk £22.50, \$39.50; pbk £9.95, \$17.95.

THIS is the fourteenth volume in a "Seminar Series" sponsored by the Society for Experimental Biology, and is the outgrowth of a one-day symposium on biological clocks held in London in early 1981. It is not, however, a typical "proceedings" volume; instead, the participants were invited after the meeting to prepare essays which could, under Brady's editorial guidance, constitute chapters in "a student text ranging widely over the whole subject".

The coverage is indeed as broad as promised. With the exception of mathematical modelling, there is at least passing mention of essentially every aspect of research on endogenous timekeeping. Four of the eleven chapters are devoted to zoological aspects of the subject; two deal only with botanical issues; one is on circadian rhythms in man; and the others cover research which knows no photosynthetic boundary. There are three essays on adaptation to season (photoperiodism); one on temporal components of animal orientation; and another which emphasizes ecological-adaptive aspects of circadian rhythms. The final chapter is an all-too-brief attempt to summarize the many recent and exciting studies, in both vertebrates and invertebrates, which have been undertaken to localize the circadian clock and to characterize the physiological mechanisms which produce the whole-animal rhythm.

The quality of the individual contributions ranges from passable to outstanding, the level of treatment from elementary to rather technical. The research interests of the contributors tend to be strongly emphasized, and although this is to some extent unavoidable, the chapter on tidal and lunar rhythms seems to me to be a particularly extreme case. In it, for example, Naylor describes research from

his own laboratory, dating from 1969 onward, on entrainment of tidal rhythms by cycles of various stimuli, and then says, "The first evidence for mechanical disturbance as an entraining agent" was published in 1975. The appropriate but uncited reference here is in fact from 1965 (*Science* 147, 864), a citation which would put Naylor's own research in a somewhat different perspective. Naylor also provides an enthusiastic discussion, with a schematic diagram, of the 1959 claim by Chapin and Wing that the Moon probably influences the breeding cycle of the wide-awake tern on Ascension Island. In fact, the relationship between Moon phase and the breeding data could scarcely be more random, but Ashmole's 1963 critique (*Ibis* 130b, 297) is unmentioned.

Author biases show in other ways as well. For example, I was surprised to read, in Saunders's chapter on photoperiodism, "It is thus possible that all photoperiodic clocks are oscillatory in nature" — all the more disconcerting since it is immediately preceded by a description of Lee's compelling and unrefuted evidence for hour-glass (non-oscillatory) timing in aphids.

But weaknesses of this sort in some of the individual chapters are a secondary issue; the more important question is whether this volume would be — as intended — a satisfactory textbook for advanced undergraduates. I think not. Overall, *Biological Timekeeping* is too disjointed and author-biased for such readers and Brady's editorial attempts to stitch together a patchwork quilt from these contributions just do not manage to give the whole an adequate sense of coherence; it remains a collection of essays. As a more coherent and definitive alternative, I would instead recommend *Biological Rhythms*, Vol.4 of the *Handbook of Behavioral Neurobiology* (edited by J. Aschoff; Plenum, 1982): it is more expensive than a paperback copy of this book, but there is much more for your money. □

J.T. Enright is Professor of Behavioral Physiology at the Scripps Institution of Oceanography, La Jolla, California.

Whether or not because of this advice, the fund seems never to have been established. Nevertheless a long succession of British and other foreign workers, including myself, have found excellent working conditions and a great intellectual stimulus at the Stazione Zoologica. We have a great deal to gain from continuing this tradition. The laboratory has been re-equipped and is anxious to receive us. The fauna is very rich and healthy, and rumours of its wholesale destruction by pollution are quite untrue, as anyone who has worked there recently will testify.

Darwin's letters contain many interesting indications of his character and capacities. He regrets that he is not a better scholar of German. He repeatedly suggests caution. He speaks often of his health. He continually shows humility and concern for the labours of others. There is not a great deal about science, though he speaks of "the cause of Wallace's sad falling away" and that "Mivart much misrepresents my views". He says of *The*

Descent of Man that "I suppose it was a mistake on my part to publish it; but anyhow it will pave the way for some better work". Indeed so it has!

Both correspondents were interested in barnacles, and at one point Dohrn suggested how the strange sac-like cirripedes are barnacles that are parasitic on crabs and other crustaceans. Darwin was for once enthusiastic: "This case seems to me the most interesting one of gradation ever recorded, viz from an animal with a stomach to one with roots like a plant".

The Stazione Zoologica and Dr Christiane Groeben the editor are to be congratulated in bringing this fascinating correspondence together. The letters and the notes tell us a lot, not only about the two chief characters but also about many others and about the state of zoology in the middle of the nineteenth century. □

J. Z. Young is Professor Emeritus of Anatomy in the University of London.

Patents for bioscience

B. J. Bate

Patenting in the Biological Sciences: A Practical Guide for Research Scientists in Biotechnology and the Pharmaceutical and Agrochemical Industries. By R.S. Crespi. Pp.211. ISBN 0-471-10151-6. (Wiley: 1982.) £16, \$38.

BY AND LARGE the world's patent systems have coped adequately with the needs of industry, and have generally provided that stimulus to innovation that is often put forward in their justification. However, the past 20 years or so have seen developments, notably in the biosciences, for which such systems were not designed.

How, for example, do you patent a complex molecule which, even if it can be obtained in reasonable purity, may be difficult to analyse, may already occur in nature and is one of a range of possibly active compounds all of which cannot economically be tested? Or, if the invention lies in the selection or even engineering of a micro-organism for use in a chemical process, how are the public — who must be put into a position to use the invention at the end of the patent monopoly period — to be given access to the organism? And how is the patentee to be protected from those who, using the techniques of genetic engineering, seek to extract the operative "bits" and introduce them into a different organism? We are only just entering the jungle of microbiologically-based inventions, and R.S. Crespi's *Patenting in the Biological Sciences* is a comprehensive yet easily digestible guide that deserves a wider readership than the title may encourage.

Following a useful analysis of the structure of the patent specification, the author discusses how different types of invention may be claimed, illustrated by reference to patents covering *inter alia* cephalosporins, vitamin B₁₂, DDT, pyrethroids, vaccines, cell lines and Chakrabarty's *Pseudomonas*. Also included is an account of the patenting of products found in nature, and methods of treatment. The chapter on genetic engineering is particularly helpful, including as it does suggestions of where patentable features may be found, and a discussion of the Budapest Treaty. For a wider readership such common problems as confidential disclosure, experimental use and inventorship are well handled.

A book that attempts so much inevitably has a few faults: the index could be more helpful; more and better charts could be used; and the importance of full disclosure of all relevant experimental results deserves a note. And, incidentally, to which of the five individuals mentioned between pages 103 and 106 — patent agent, patent adviser, draftsman, patent attorney or practitioner — should the researcher look for advice? □

B. J. Bate is Manager of the Patents Department of the Heavy Chemicals Business Area of ICI.

The psychology of babyhood

George Butterworth

Development in Infancy, 2nd Edn. By T.G.R. Bower. Pp.304. Hbk ISBN 0-7167-1301-2; pbk ISBN 0-7167-1302-0. (W.H. Freeman: 1982.) Hbk £15.80, \$20; pbk £7.40, \$9.95.

TOM Bower has long held a reputation as the *enfant terrible* of baby psychology. An otherwise sympathetic reviewer nevertheless said of the first edition of this seminal book that "it would serve for years as a source of ideas for doctoral theses that will prove it inconclusive theoretically and wrong or partially wrong experimentally" (J.S. Bruner *Nature* 252, 514; 1974). Perhaps in response, Bower prefaces the second edition with the remark that some of his more outrageous speculations now seem positively conservative! This revision, extended by 50 pages, will both delight and infuriate for it retains all the originality and virtues — and some of the faults — of its predecessor.

Infancy researchers world-wide would now admit that the experimental evidence of the past eight years has vindicated many of Bower's claims. About half of his new material consists of reviews of experiments that corroborate his own studies of infant perception. For example, certain types of intersensory coordination such as that between seeing and hearing; perception of size and shape constancy or aspects of visual manual coordination are innately attuned to the characteristics of the environment.

Bower adds a new chapter on social development and this is logical, for what better purpose could an infant's precocious

perceptual competence serve than to adapt the baby to the social world? Here he discusses a tantalizing new finding of gender identification among babies in their first year. It seems that infants, given the choice, prefer to observe a motion film of a baby of the same sex; they recognize the similarity between the other boy or girl and themselves, perhaps through sex-linked patterns of movement originating in differences in skeletal structure. A problem is that further details of this interesting phenomenon may prove hard to come by because the reference cited is unpublished.

For the most part, however, Bower is careful to qualify his claims in this new edition and he draws attention to the methodological problems involved in observing such unexpected phenomena in young babies.

The first edition of *Development in Infancy* had the effect of stimulating an original kind of research on the capabilities of babies. It has helped to usher in an ecological approach to problems of perception and cognition, at least in developmental psychology, which makes the remarkable capacities of young babies so much easier to understand. This new edition should inspire readers to examine further the functional significance of these capacities in laying the foundations for human social behaviour. □

George Butterworth is Senior Lecturer in Psychology at the University of Southampton. He is editor of Infancy and Epistemology: An Evaluation of Piaget's Theory (Harvester, 1982).